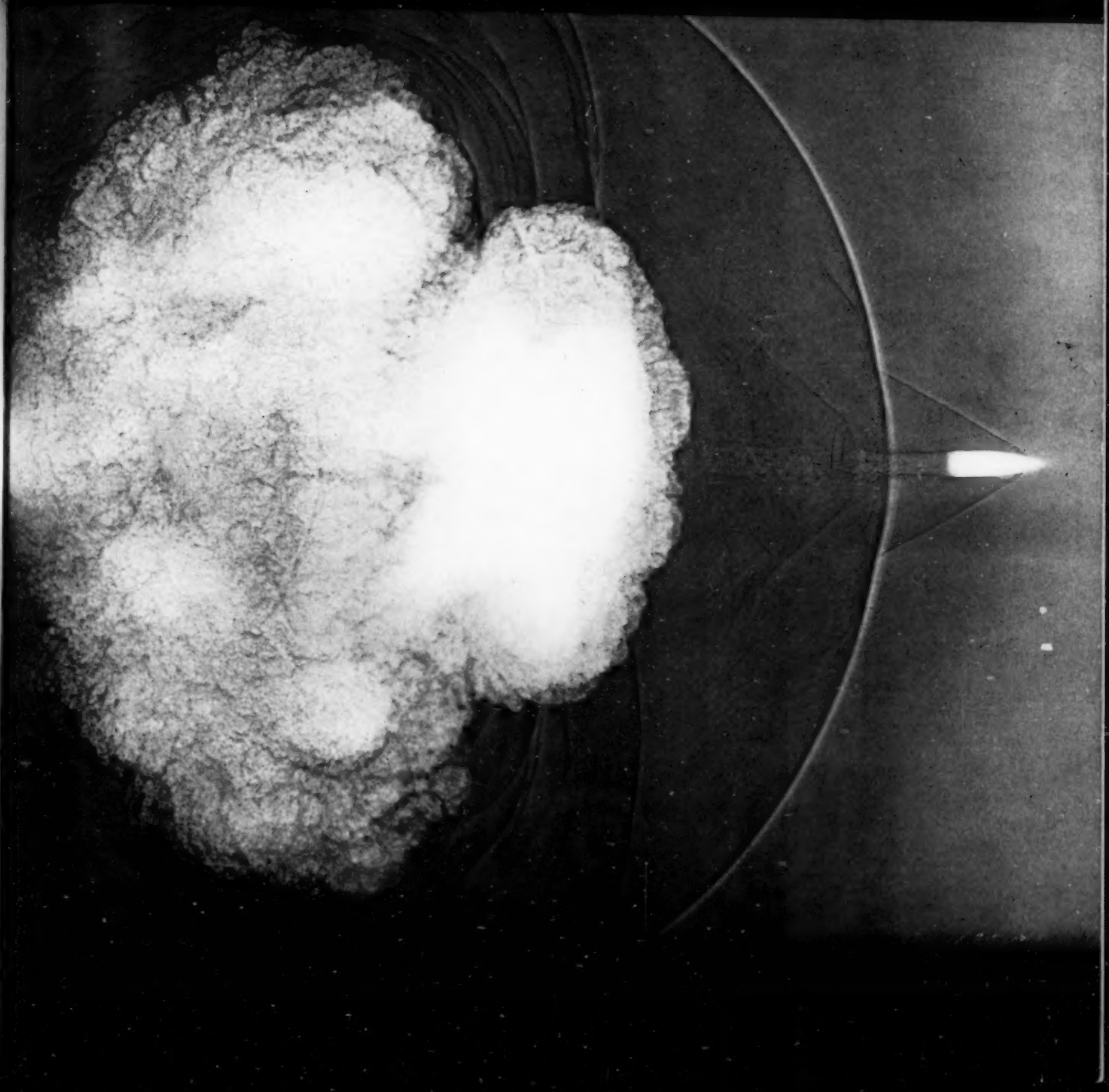
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 21, No. 5 MAY, 1950

In This Issue

	Page
On the Shielding Qualities of Different Concrete Mixtures P. C. Gugelot and M. G. White	369
Bending Method for Measurement of the Tensile Strength and Young's Modulus of Glass Fibers David Sinclair	380
Magnetic Field Calculations for Large Cross-Section Cloud-Chamber Coils H. W. Koch	387
Effect of Compressive Stresses on the Linear Thermal Expansion of Magnesium and Steel Joseph L. Rosenholtz and Dudley T. Smith	396
Thermal Noise at High Frequencies A. van der Ziel	399
Semiconductive Colloidal Suspensions with Non-Linear Properties Hans E. Hollmann	402
Internal Friction of Metals at Very High Temperatures T'ing-Sui Kê	414
Orientation Dependence of the Rate of Grain Boundary Migration Paul A. Beck, Philip R. Sperry, and Hsun Hu	420
Measurement of the Thermal Diffusivity of a Soil by the Use of a Heat Pump E. B. Penrod	425
I. Indoor Measurement of Microwave Antenna Radiation Patterns by Means of a Metal Lens G. A. Woonton, R. B. Borts, and J. A. Carruthers	428
Load Induced X-Ray Line Broadening in Nylon Filaments John Kauffman and Waller George	431
Aluminum Oxide Replicas for Electron Microscopy Produced by a Two-Step Process G. Hass and M. E. McFarland	435
Diffusional Viscosity of a Polycrystalline Solid Conyers Herring	437
Streak Photography Irwin Vigness and R. C. Nowak	445
Surface Tension Effects in Oriented Thin Silver Films G. W. Johnson	449
Upper and Lower Bounds of Eigenvalues for Composite-Type Regions E. L. Chu	454
Letters to the Editor:	
Single Crystal X-Ray Techniques for Use with Alloys Requiring Protective Atmospheres Joseph Singer and F. H. Ellinger	467
Errata: On the Diffraction of an Electromagnetic Wave through a Plane Screen J. W. Miles	468
Erratum: Distribution of Phases in Two-Phase Solids Leonard D. Jaffe	468
New Books	468
Index to Advertisers	xxviii

BOARD OF EDITORS
Elmer Hutchisson, *Editor*

Associate Editors
Howard H. Aiken
Otto Beeck
James Hillier
W. James Lyons
Philip Morrison
Eric Reissner
Raymond J. Seeger
Frederick Seitz
Charles S. Smith

Publication of the American Institute of Physics

OFFICERS
George R. Harrison, *Chairman*
George B. Pegram, *Treasurer*
John T. Tate, *Adviser on Publications*

GOVERNING BOARD
J. W. Beams
C. Paul Boner
Wallace R. Brode
W. F. Fair, Jr.
J. B. Fisk
G. P. Harnwell
George R. Harrison
W. V. Houston
Rudolph Kingslake
Paul Kirkpatrick
Paul E. Klopsteg
F. W. Loomis
Philip M. Morse
J. R. Oppenheimer
George B. Pegram
Harold K. Schilling
Francis G. Slack
John C. Steinberg
John T. Tate

Consultant on Printing
Melvin Loos

ADMINISTRATIVE STAFF
Henry A. Barton, *Director*
Wallace Waterfall, *Executive Secretary*
Theodore Vorburger, *Assistant Treasurer and Comptroller*
Ruth F. Bryans, *Publication Manager*
Betty Hofstadter, *Assistant Secretary*
Eileen B. Neuberger, *Circulation Manager*

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, not submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members and associate members of the American Institute of Physics	$7.00	$ 8.00
To all others	9.00	10.00

Back Number Prices

Yearly back number rate when complete year is available: $10.00. Single copies: $1.00 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Copyright 1950, by the American Institute of Physics. Permission is hereby given to reproduce freely author abstracts in abstract journals.

Journal
of
Applied Physics

Volume 21, Number 5 May, 1950

On the Shielding Qualities of Different Concrete Mixtures*

P. C. Gugelot and M. G. White
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received April 4, 1949)

Absorption curves for neutrons and γ-radiation produced by 16-Mev protons on beryllium are measured in absorber blocks of various concrete mixtures to obtain information on their shielding properties. It is shown that a mixture of cement, scrap iron, and limonite ore is as efficient as water in slowing down and capturing neutrons and, in addition, has sufficient density ($\rho=3.5$ g/cm^3) to be an effective absorber for γ-radiation. Compared to ordinary concrete, a three-foot wall of iron and limonite concrete is superior by a factor of 280 for neutrons and by a factor of 20 for γ-radiation. Boron is shown to be of little value in this type of shielding. Mechanical properties of the scrap iron-limonite concrete are satisfactory for structural purposes. From the practical cost standpoint, a concrete made entirely of limonite ore is shown to be a reasonable compromise.

I. INTRODUCTION

With the construction of larger accelerators and nuclear reactors, the problem of shielding against radiation becomes more and more important. It is not only necessary to protect against health hazards, but also to reduce the background for experiments.

Water shields commonly used around cyclotrons are effective in slowing down and capturing neutrons of a few Mev. However, for very high neutron energies the slowing down qualities of this medium decreases, owing to the small scattering cross section of fast neutrons in hydrogen. In addition, the absorption of γ-radiation in water is negligible. It is, therefore, necessary to find a shielding material which is effective in slowing down neutrons, and with sufficient density to be an efficient γ-ray absorber.

The theory of slowing down of neutrons has been worked out in a hydrogenous medium by Fermi.[1] His theory describes the mean square distance $(r^2)_{Av}$ which a neutron with the initial energy E_0 travels in being slowed down to an energy E_1. The "Age-theory," developed by Bethe, Korf, and Placzek,[2,3] calculates the spatial distribution of neutrons in a medium. By using the age of a neutron:

$$\tau=1/6(r^2)_{Av}=\frac{1}{3[1-(2/3A)]\xi}\int_{E_1}^{E_0}\lambda^2\frac{dE}{E}$$

$$\xi=1-\frac{A}{2}[1-(1/A)]^2\log(A+1)/(A-1),$$

where A = atomic weight of element of which the absorber is composed, and λ = mean free path of the neutrons of energy E, the slowing down equation can be written in a simple form:

$$\Delta q=\partial q/\partial\tau.$$

q = number of neutrons per cm^3 slowed down to energy E_1, the solutions of which are Gaussian functions. Placzek[4] and Marshak[5] extended the theory to absorbers composed of heavy nuclei, and to mixtures of different nuclei. Higher approximations were developed showing that the Gaussian distribution of neutrons of final energy E_1, obtained by the age-theory, is valid only to distances equal to τ/λ and then only when many collisions are necessary to slow the neutrons down ($A\gg1$). At larger absorber thickness the neutron density decreases exponentially. These calculations are all based on slowing down of neutrons by elastic collisions;

* Assisted by the joint program of the ONR and the AEC.
[1] E. Fermi, Ricerca Scient. **7**, 13 (1936); H. A. Bethe, Rev. Mod. Phys. **9**, 69 (1937).
[2] Bethe, Korf, and Placzek, Phys. Rev. **57**, 573 (1940).
[3] E. Fermi, Los Alamos Lectures (February, 1946).
[4] G. Placzek, Phys. Rev. **69**, 423 (1946).
[5] R. E. Marshak, Rev. Mod. Phys. **19**, 185 (1947).

therefore, they are only valid for initial neutron energies lower than the first excited states in nuclei. The poor knowledge about nuclear excitation levels makes calculations above these energies uncertain and experiments are needed.

Measurements were carried out by Munn and Pontecorvo[6] with heterogeneous mixtures of water with iron or bismuth. They obtained the mean square distance $(r^2)_{Av}$ required to slow Ra+Be neutrons down to 1.44 ev (In detectors) and to thermal energy (Dy detectors). Values observed by these authors are:

for water: $(r^2)_{Av\text{-}In}=278$ cm^2, $(r^2)_{Av\text{-}th}=330$ cm^2,
for water+Bi (density 5.3 g/cm^3): $(r^2)_{Av\text{-}th}=562$ cm^2,
for water+Fe (density 3.9 g/cm^3): $(r^2)_{Av\text{-}In}=319$ cm^2,

showing that it is possible to use mixtures of water and iron in which $(r^2)_{Av}$ is nearly the same as in pure water, but with the greatly improved γ-ray absorption of the iron.

Tittle and Goodman,[7] in order to carry out absorption experiments in which the components of the mixture are easily interchangeable, studied the difference in neutron distribution between homogeneous and heterogeneous mixtures of C, O, and H. Neutrons from a Ra+Be source were used and In-resonance neutrons detected. The intensity distribution was found to be the same for the different mixtures with equal C, O, and H concentrations at distances >20 cm from the source.

Hildebrand and Moyer[8] obtained absorption coefficients for 80-Mev neutrons detected by an ionization chamber. The half-value thickness for water is 55.9 cm and for concrete (ordinary and barite) 24.1 cm.

Concrete with 5–10 percent water was also investigated by Jensen and Ritter[9] with neutrons from a Ra+Be source. Spherical and plane geometry were used. In the first case, a distribution for In-resonance neutrons was obtained of the form $r^2I=$ const. $\exp(-r/L')$ ($r>40$ cm), $L'=21.8$ cm. For the plane geometry the intensity decreased exponentially: $I=$ const. $\exp(-r/L)$ ($r>40$ cm), $L=15.3$ cm. The authors show that, at absorber thickness where the fast neutron flux is in equilibrium with the slow neutron flux for spherical geometry, $Ir=$ const. $\exp(-r/L_s)$, in which $L_s=L$, and $1/L'=1/L-1/r$. A water wall of 65-cm thickness reduces the neutron flux in this case as much as 160 cm concrete ($I/I_0=10^{-3}$).

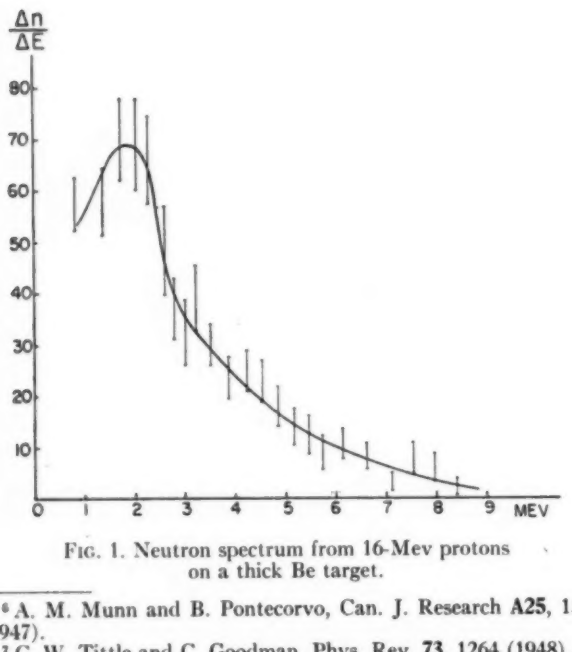


FIG. 1. Neutron spectrum from 16-Mev protons on a thick Be target.

II. EXPERIMENTAL ARRANGEMENT

Before the radiation shield around the Princeton frequency modulated cyclotron[10] was constructed, we made absorption measurements with different concrete mixtures in order to design a wall which would reduce background sufficiently to make possible careful experiments with the deflected proton beam. Cubic blocks of 46-cm edge were cast, having a steel tube of 7.5-cm diameter through the center of the block. Slabs of 5.3-cm thickness, prepared from the same mixture, fitted smoothly into the tube. In this way slow neutron detectors or a Geiger-Müller counter could easily be placed between the slabs and an almost homogeneous absorber was obtained. The absorber blocks were placed in a reproducible position in the forward direction of neutrons generated by 16-Mev protons on a thick Be target. To obtain information on the energy of the neutrons and to be able to compare our results with the experiments which were carried out with a Ra+Be neutron source, we studied the neutron spectrum by measuring proton-recoil tracks in 100μ Ilford C2 photographic emulsions (Fig. 1). Tracks in the forward direction, originating and ending in the emulsion, were measured and corrected for scattering cross section[11] and for loss of tracks in a layer of finite thickness. The plate was bombarded outside the acceleration chamber 52 cm from the internal Be target. An appreciable number of the low energy tracks is due probably to neutrons which suffered collisions in the iron of the magnet. This spectrum differs from a Ra+Be neutron spectrum[12] in the low energy range; however, the high energy tail is in both cases somewhat similar. In a total of 900 tracks, five were found with energy >9 Mev.

To be effective, a shield must strongly absorb fast neutrons, slow neutrons and gamma-rays since all these particles are radiated by nuclear machines and also result from the stopping of fast neutrons. A complete

[6] A. M. Munn and B. Pontecorvo, Can. J. Research **A25**, 157 (1947).
[7] C. W. Tittle and C. Goodman, Phys. Rev. **73**, 1264 (1948).
[8] R. Hildebrand and B. J. Moyer, Phys. Rev. **72**, 1258 (1947).
[9] P. Jensen and O. Ritter, Zeits. f. Naturforschung **2a**, 376 (1947).
[10] White, Fulbright, Gugelot, and Bush, Phys. Rev. **74**, 1242A (1948).
[11] Goldsmith, Ipsen, and Feld, "Neutron cross section of the elements," Rev. Mod. Phys. **19**, 259 (1947); D. Bohm and C. Richman, Phys. Rev. **71**, 567 (1947).
[12] D. L. Hill, Argonne National Lab., Declassified Report No. CP-3800.

study of shielding materials should include attenuation measurements on all these types of incident radiation using detectors which are specific for the radiation in question. However, for neutron studies, we have confined ourselves to the use of thermal or silver resonance neutron detectors and for gamma-rays a copper walled Geiger-Müller counter was used. The attenuation of fast neutrons we infer from the slopes of the slow neutron intensity curves at large absorber thickness since all *incident* slow neutrons are removed by the early part of the absorber and only fast neutrons are left to produce the detected slow neutrons. Considerable care must be exercised in using data obtained near the back of a block because room-scattered low energy neutrons can falsify the curves.

The slow neutron intensity in the absorber blocks due to the total incident cyclotron flux was obtained by activating silver foils of 2.5-cm diameter and 0.7-mm thickness between the slabs and the intensity of the 2.37^m Ag^{108}-period was measured with a thin wall G-M tube. The counting rate was corrected for decay and background and then compared with the activity of a monitor foil. To obtain the activation by resonance neutrons, the Ag foils were placed in an envelope of 0.45 g/cm^2 Cd (Ag^{108} has resonance levels at 50 and 15 ev). Silver was chosen because of its convenient half-life. The size of the punched foils was identical. In general, only the front side of the foils was measured. All these data are given in the "*B*" curves of Figs. 3–10. Runs in which the activity on both sides was investigated showed that the error introduced by measuring only one side was less than 4 percent. An uncertainty in foil position of 1–2 mm resulted from the rather rough faces of the slabs.

The diffusion of *thermal* neutrons incident on the test blocks was studied by interposing water cans between the cyclotron and test blocks and then taking attenuation curves with and without cadmium absorber placed *between* the water cans and the test block. These data are the "*A*" curves in Figs. 3–10. According to the diffusion theory, for plane geometry the intensity of thermal neutrons decreases exponentially. The $1/e$-value thickness is the "diffusion-length" l; it is connected with the transport mean free path λ and capture mean free path Λ by $l=(\frac{1}{3}\lambda\Lambda)^{\frac{1}{2}}$.

The influence of the finite size of the test blocks on the absorption curves was estimated in the following way: The neutron attenuation curve obtained in the cement +iron+limonite ore test block was compared with the attenuation curve obtained by placing water cans in front of the test block leaving an opening of 15×15 cm in the center. This in effect decreases the size of the absorbing block. That this is the actual case is shown by the slow neutron intensity distribution perpendicular to the fast neutron beam (Fig. 2). The neutron density obtained with the water can "diaphragm" (curve 2) has a pronounced maximum in the center, and the slope of the attenuation curve was unchanged from that

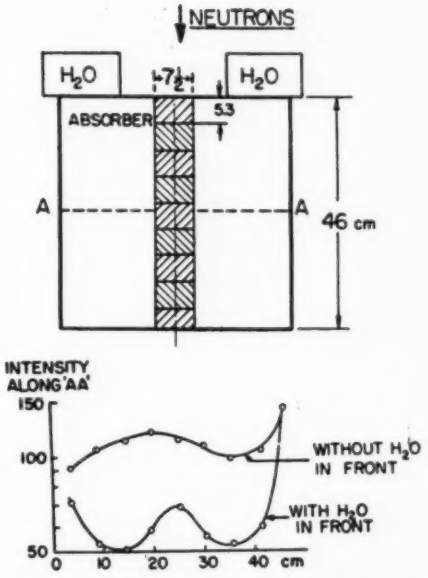


FIG. 2. Absorber block and thermal neutron distribution inside the absorber perpendicular to the fast neutron flux, showing the effect of neutron collimation by water tanks in front of the absorber.

obtained when the entire block was irradiated. The slow neutron intensity at large thickness of absorber is mainly caused by neutrons slowed down by a large number of small angle deflections, and the subsequent diffusion of slow neutrons contributes very little to the observed half-value thickness of the attenuation curve.

Gamma-ray attenuation, when measured by exposing test blocks to the entire cyclotron radiation and studying the counting rate of a copper walled Geiger-Müller counter cannot be expected to yield precise results since a few counts will be due to fast neutrons causing knock-on ions in the counter. More important, the fast neutrons will create gamma-rays in the test block. Nevertheless, as a practical matter, the attenuation of a shield as measured by a Geiger-Müller counter is of considerable interest. Figure 11 collects all the Geiger-Müller counter data for the various absorbers.

III. RESULTS FROM NEUTRON ABSORPTION EXPERIMENTS

a. Water

For this experiment a tank was used of larger dimensions than the concrete absorbers. The neutron detectors were placed in a hard rubber holder. Figure 3 shows the absorption curves so obtained. Figure 3A gives the attenuation of *thermal* neutrons which strike the water tank after being generated in a preceding layer of water. Curve 1 corresponds to the intensity distribution resulting from *all* incoming neutrons, curve 2 to the intensity distribution resulting from the faster than thermal neutrons which succeed in passing a cadmium sheet placed between the test water tank and the slow neutron generating water cans. The difference of these two curves shows the exponential decrease in

the intensity of incoming thermal neutrons. The diffusion-length is $l=2.7$ cm; it compares within the error to the value 2.88 cm obtained for water by Allison *et al.*[13] with In detectors. The deviation is caused by the uncertainty in position of the Ag-foils and by the Cd shield which eliminates neutrons reflected by the water cans back into the tank. Curve 3, of Fig. 3B, shows the total Ag activity caused by exposing the test tank to the direct cyclotron radiation, curve 4 the intensity produced by resonance neutrons only and the difference of 3 and 4 corresponds to the thermal neutron intensity. This curve yields a half-value thickness of 5.5 cm for absorber thickness larger than $r=20$ cm. From 3B it is possible to calculate $(r^2)_{Av}$ for thermal and resonance neutrons:

$$(r^2)_{Av}=\frac{\int_0^\infty I(r)r^2dr}{\int_0^\infty I(r)dr}$$

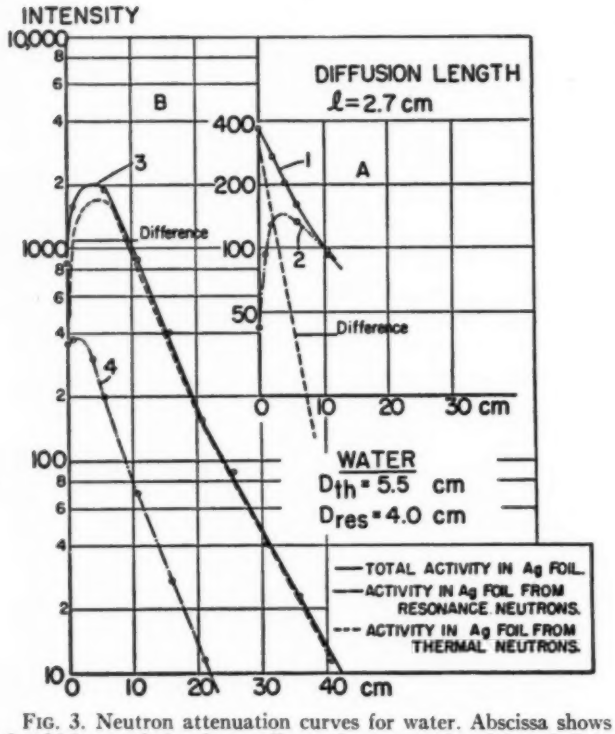


FIG. 3. Neutron attenuation curves for water. Abscissa shows the thickness of absorber, ordinate the corrected activity of silver detectors. A. Attenuation curves for incoming thermal neutrons, a layer of 15-cm water is placed between absorber and cyclotron. Curve 1. No Cd between absorber and waterlayer. All neutrons after passing through the 15-cm thick waterlayer can enter the absorber. Curve 2. Cd between absorber and waterlayer. Only faster than thermal neutrons enter the absorber tank. The difference of 1 and 2 shows the attenuation of incoming thermal neutrons from which the diffusion length of thermal neutrons is calculated. B. Attenuation curve of incoming fast neutrons. The absorber is only shielded by Cd. Curve 3. Ag detectors sensitive to all neutron energies. Curve 4. Ag detectors wrapped in Cd and thus are only sensitive to faster then thermal neutrons. The difference between 3 and 4 shows the Ag activity from thermal neutrons. From this curve the half-value thickness D_{th} is obtained.

[13] See J. H. Rush, Phys. Rev. **73**, 271 (1948).

$I(r)$: The activity of the detectors, which is proportional to the slow neutron-intensity at position r.

For thermal neutrons: $(r^2)_{Av\text{-}th}=119.2$ cm^2.

For resonance neutrons: $(r^2)_{Av\text{-}res}=62.3$ cm^2.

For the Ag detector no precise neutron energy is connected with $(r^2)_{Av\text{-}res}$. The corresponding migration area is for thermal neutrons:

$$M_{th}^2=1/6(r^2)_{Av\text{-}th}=19.9 \text{ cm}^2.$$

The age for resonance neutrons is:

$$\tau_{res}=1/6(r^2)_{Av\text{-}res}=10.4 \text{ cm}^2.$$

According to the diffusion theory we have: $M_{th}^2=\tau_{th}+l^2$. The value for l is 2.88 cm. Hence the age of thermal neutrons is $\tau_{th}=11.8$ cm^2. The difference in age between thermal and resonance neutrons is: $\tau_{th}-\tau_{res}=1.3$ cm^2. This value is of the expected order and would correspond to an average resonance energy of about 30 ev. Comparing the age of our thermal neutrons with τ_{th} for thermal neutrons from a Ra+Be source slowed down in water, for which Anderson *et al.*[14] obtained 46.4 cm^2, shows that we have to deal with a neutron spectrum which extends to very low energies. It was indeed necessary to shield the front of the absorber with Cd to obtain useful values for $(r^2)_{Av\text{-}th}$.

b. Water+Boron

It has frequently been suggested that water shields could be improved by adding boron in the form of borax. Two beneficial effects are to be expected. First,

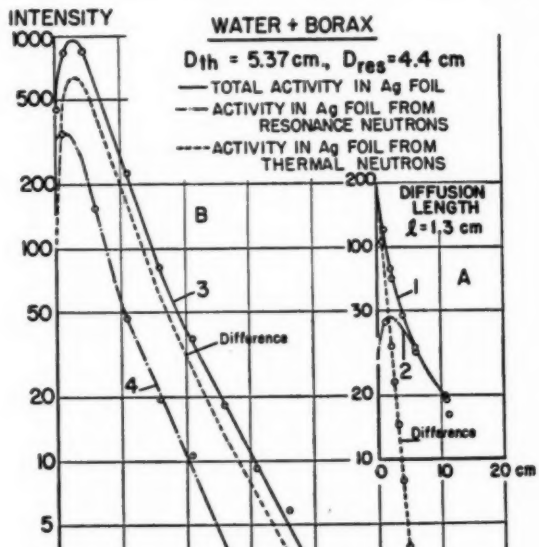


FIG. 4. Neutron attenuation curves for water+3 percent borax.

[14] See J. H. Rush, Phys. Rev. **73**, 271 (1948).

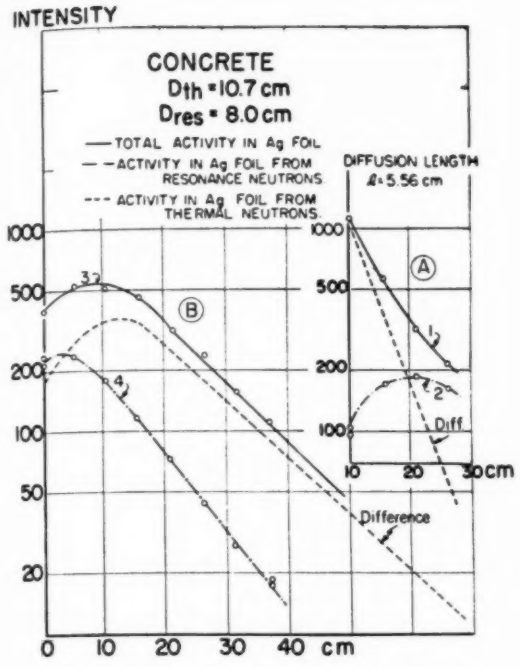


FIG. 5. Neutron attenuation curves for concrete.

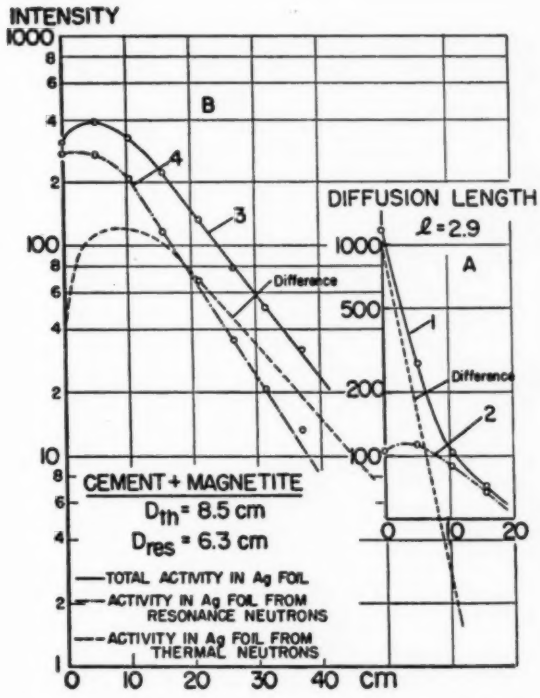

INTENSITY
DIFFUSION LENGTH
ℓ = 2.9
Difference
CEMENT + MAGNETITE
Dth = 8.5 cm
Dres = 6.3 cm
— TOTAL ACTIVITY IN Ag FOIL
— – ACTIVITY IN Ag FOIL FROM RESONANCE NEUTRONS
--- ACTIVITY IN Ag FOIL FROM THERMAL NEUTRONS
0 10 20 30 40 cm

FIG. 6. Neutron attenuation curves for concrete+Pyrex.

the boron capture of neutrons tends to suppress the neutron-proton capture gamma-rays and second, the large boron cross section should reduce the diffusion of slow neutrons. To test the latter point a 3 percent solution of borax in water was substituted for the previously studied pure water. The capture mean free path Λ can be calculated assuming $\sigma H_a = 0.4\times10^{-24}$ cm^2 [15,16] and $\sigma B_a = 700\times10^{-24}$ cm^2. We obtain

$$\frac{\Lambda \text{ solution}}{\Lambda \text{ water}} = \frac{1}{6.2}.$$

From this it follows that

$$\frac{l \text{ solution}}{l \text{ water}} = \left(\frac{\Lambda \text{ solution}}{\Lambda \text{ water}}\right)^{\frac{1}{3}} = \frac{1}{2.5}.$$

By experiment (Fig. 4A) we observe l solution $= 1.3$ and for pure water l water $= 2.85$ (Fig. 3A) yielding the ratio

$$\left(\frac{l \text{ solution}}{l \text{ water}}\right)_{\text{exp}} = \frac{1}{2.2}.$$

The agreement between the predicted and observed ratio is reasonably good. However, the half-value thickness obtained from Fig. 4B, which includes *all* neutrons both fast and slow incident upon the borax solution absorber is the same as found for pure water showing that the slope of the attenuation curves is mainly determined by the larger mean free path of the fast neutrons present in the cyclotron radiation.

[15] Manley, Haworth, and Luebke, Phys. Rev. **61**, 152 (1942).
[16] E. J. Harringh and J. L. Stewart, Can. J. Research **19**, 33 (1941).

It is of interest to compare the effects of boron absorption on the flux of thermal neutrons and silver resonance neutrons. The area under an absorption curve, such as curve 3 Fig. 3B, is *proportional* to the integral

$$F_E = \int_0^\infty I_E(r)dr$$

where $I_E(r)$ is the number of neutrons at thickness r which have been slowed down to the energy E at which the detector is efficient. Thus if there is strong absorption at an energy *higher* than E the integral will be small. By extrapolating the curves of Figs. 3B and 4B we obtain for resonance energy neutrons the following data: In water $F_{res} = 2.7\times10^3$; in water plus borax $F_{res} = 2.3\times10^3$ showing that in spite of the boron the density of Ag. *resonance* neutrons remain unchanged. However, for *thermal* neutrons $F_{th} = 2.1\times10^4$ in water, and $F_{th} = 5.6\times10^3$ in borax solution indicating that the boron has decreased the number of neutrons in the thermal energy range.

c. Concrete and Concrete+Pyrex

A sample of ordinary concrete was tested in the manner described above. The results are given in Fig. 5. The diffusion length for thermal neutrons is $l = 5.6$ cm. The half-value thickness for the neutrons slowed down to thermal velocity is $D_{th} = 10.7$ cm. This mixture contains 0.21 g water/cm^3, a number obtained from the amount of water used to make the mixture and from the moisture in the gravel and sand. (A small amount of

water will have evaporated during the setting of the concrete.) Composition: 46.3 percent gravel by weight, 27.6 percent sand, 17.7 percent cement and 8.4 percent water; density $\rho=2.35$ g/cm³. The large half-value thickness is due to the large number of collisions with relatively heavy nuclei which neutrons have to suffer in being slowed down; inelastic scattering will be negligible for Si and O.

Another sample was made containing 1.25 percent Pyrex glass by weight, 46.55 percent gravel, 27.78 percent sand, 18.37 percent cement and 6.05 percent water. The boron had an observable effect on the diffusion length (Fig. 6): $l=3.3$ cm (Pyrex contains 3 percent boron). The ratio of hydrogen atoms to boron atoms is in this case 2.7×10^2. If boron and hydrogen were solely responsible for capturing thermal neutrons the ratio of the diffusion lengths would be, taking into account the different hydrogen contents in both mixtures,

$$\frac{l\ \text{concrete+Pyrex}}{l\ \text{concrete}}=\frac{1}{2.7};$$

observed is

$$\left(\frac{l\ \text{concrete+Pyrex}}{l\ \text{concrete}}\right)_{\text{exp}}=\frac{1}{1.7}.$$

This deviation is to be expected because the thermal neutron capture cross sections of Si ($\sigma_a=0.25\times10^{-24}$ cm²) and of other elements of which the absorber is composed were not taken into account. The half-value thickness of the fast neutron attenuation curve decreased to $D_{th}=7.4$ cm. It is not clear whether the Pyrex had any function in decreasing the absorption coefficient D_{th}. This mixture contains only 0.14 g water/cm³. (Density $\rho=2.39$ g/cm³.)

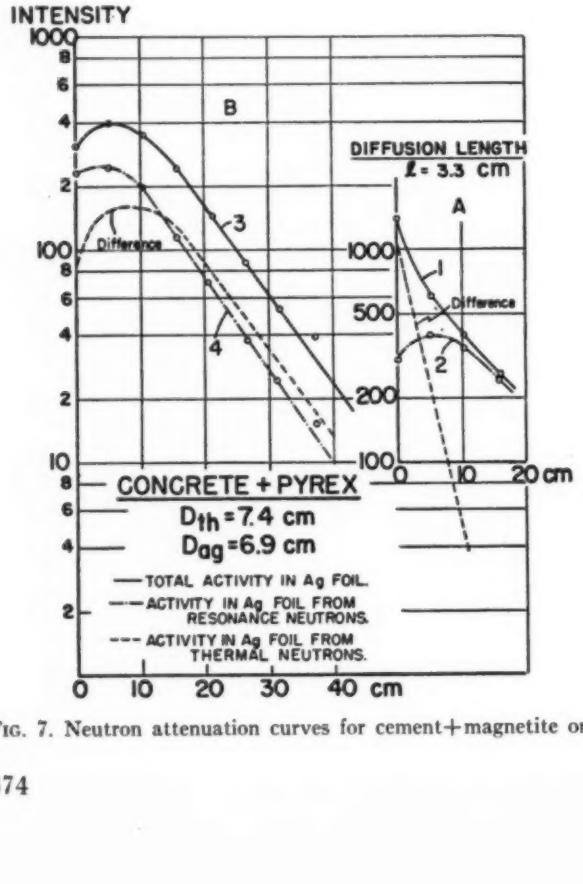


FIG. 7. Neutron attenuation curves for cement+magnetite ore.

d. Cement+Magnetite

No improvement was obtained by replacing the sand and gravel by magnetite ore (Fe_3O_4). This mixture contains 83.5 percent magnetite ore, 11 percent cement, and 5.5 percent water by weight. This amounts to 0.21 g water/cm³ and 1.93 g Fe/cm³; the density is 3.78 g/cm³. The half-value thickness is 8.5 cm (Fig. 7), slightly better than for standard concrete. Because of the iron used in this mixture the diffusion length for thermal neutrons became 2.9 cm. Fe has a relatively large capture cross section for thermal neutrons, $\sigma_{Fe_a}=2.05\times10^{-24}$ cm².[17]

e. Cement+Limonite, Cement+Limonite+Iron, Cement+Limonite+Iron+Pyrex

It was possible to increase the amount of hydrogen in these mixtures by using limonite ore ($2Fe_2O_3\cdot3H_2O$).*

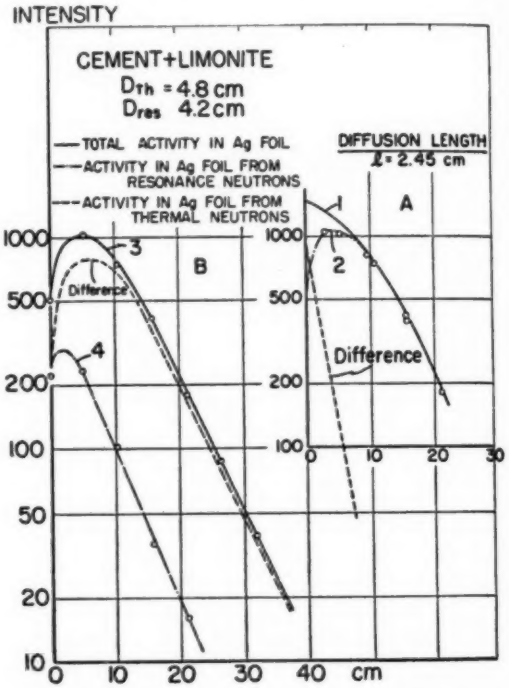


FIG. 8. Neutron attenuation curves for cement and limonite ore

The first absorber contains 68.5 percent limonite, 21 percent cement and 10.5 percent water. Taking the amount of moisture and water of crystallization into account 0.49 g water/cm³ is obtained and 0.93 g Fe/cm³. The density is $\rho=2.63$ g/cm³. The smallest half-value thickness D_{th} was obtained with this absorber: $D_{th}=4.8$

[17] J. W. Coltman and M. Goldhaber, Phys. Rev. 69, 411 (1946).

* Obtained from the Shook-Fletcher Company, Birmingham, Alabama.

cm (Fig. 8). However, the density was not sufficient for satisfactory γ-ray absorption (Section IV). Scrap iron in the form of punchings (about 2 cm diameter and about 1 cm thick) was mixed into the following block (Fig. 9): scrap iron 59.4 percent, limonite 26.35 percent, cement 9.54 percent and water 4.71 percent. A density of 4.41 g/cm³ was obtained at the expense of lower hydrogen content: 0.35 g water/cm³ and 3.23 g Fe/cm³. The half-value thickness D_{th} increased to $D_{th}=5.4$ cm.

For the last block a mixture was obtained with less iron and more hydrogen than in the former absorber. The capture probability for thermal neutrons was again increased by the use of Pyrex: scrap iron 41.4 percent,

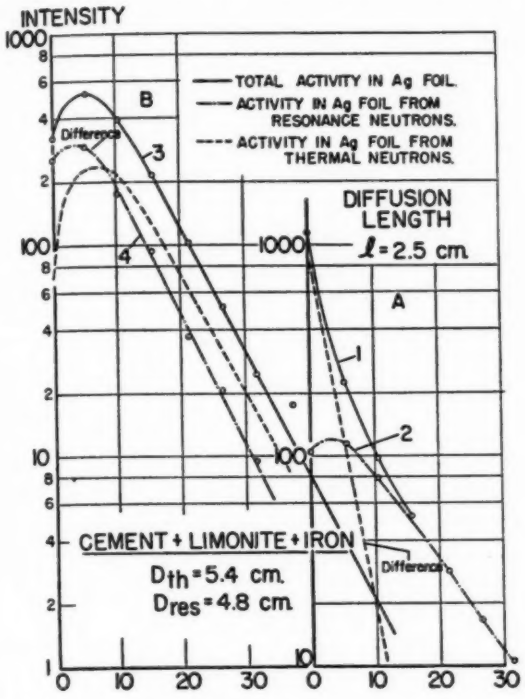


FIG. 9. Neutron attenuation curves for cement +limonite+scrap iron.

limonite 32.8 percent, cement 12.3 percent, Pyrex 0.7 percent, and water 12.8 percent. The density is 3.6 g/cm³. This amounts to: 0.45 g water/cm³ and 2.26 g Fe/cm³. The half-value thickness obtained for thermal neutrons is $D_{th}=5.0$ cm. This smaller D_{th} is probably caused by the larger amount of water per cm³. Because of the Pyrex used in this absorber, the diffusion length for thermal neutrons became $l=2.3$ cm, instead of 2.5 cm for the former two blocks (Fig. 10). In the cement +iron+limonite+Pyrex block the capture mean free path is, for hydrogen, 83 cm; for iron, 20 cm; for boron, 35 cm.

IV. γ-RADIATION

A Geiger Müller counter was placed between plugs in the test block to obtain information on the background which is to be expected outside the radiation shield when working with counters. A mica window "Victoreen" counter was used with the window closed by a brass cover. The counting rate is also connected with the neutron flux in the absorber because of (n, γ) and $(n, n\gamma)$ processes near the counter and (n, p) and (n, α) reactions in the counter wall. Therefore, with a bad neutron absorber, no steep absorption curves for the counter can be obtained. Figure 11 shows the results for all investigations. No simple exponential relation exists between intensity and absorber thickness. The largest half-value thickness, D_C, was observed for water. D_C remained unchanged in the boron solution, because only multiplication of the intensity by a constant factor <1 is to be expected for the neutron-capture γ-radiation. (All curves are normalized at $r=0$.) 1.25 cm of lead around the counter decreased the counting rate in water to 1/3.4 of its value at $r=0$ and to 1/4.1 at $r=30$ cm. More information is needed about the γ-ray spectrum and the fraction of counts caused by γ-radiation if one is to make detailed calculations.

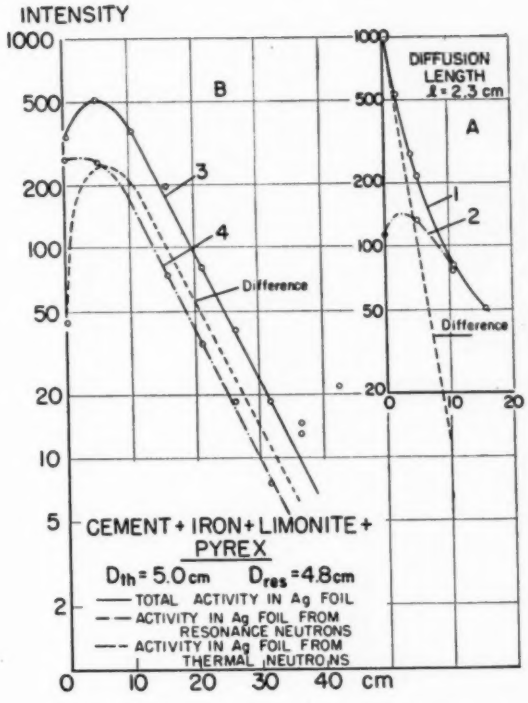


FIG. 10. Neutron attenuation curves for cement +limonite+scrap iron+Pyrex.

For both cement+limonite and cement+magnetite mixtures the same curves were observed. In the more dense magnetite the γ-ray absorption is probably higher than in the former, but this effect is obliterated by low neutron absorption in the magnetite block.

Almost exponential decays with half-value thickness of 7.9 cm are obtained in the limonite+iron absorbers.

V. DISCUSSION

By placing a water shield between an absorber block and the cyclotron, we noticed that the half-value thick-

TABLE I. Summarized data from absorption measurements.

Absorber	g H_2O/cm^3	g Fe/cm^3	Density g/cm^3	½ intensity thickness cm	Neutrons "Age" (cm^2) of thermal neutrons	Diffusion length (cm)	G-M counting rate ½ intensity thickness at $d=20$ cm	Transmission factor for a 90-cm thick wall	
								Neutrons	Counting rate
H_2O	1	—	1.00	5.5	11.8	2.7	14.8	$9\cdot10^{-6}$	$3\cdot10^{-2}$
H_2O+3% Borax	1	—	1.03	5.4	12.0	1.3	14.8	$6\cdot10^{-6}$	$3\cdot10^{-2}$
Concrete	0.2	—	2.35	10.7	86	5.6	13.3	$4.5\cdot10^{-3}$	$1\cdot10^{-2}$
Concrete+1.2% Pyrex	0.14	—	2.39	7.4	49	3.3	13.3	$3.8\cdot10^{-4}$	$7\cdot10^{-3}$
Cement+magnetite ore	0.21	1.93	3.78	8.5	70	2.9	9.0	$1\cdot10^{-4}$	$2\cdot10^{-3}$
Cement+limonite ore	0.49	0.93	2.63	4.8	23.7	2.5	9.0	$4.8\cdot10^{-6}$	$2\cdot10^{-3}$
Cement+Fe+limonite ore	0.35	3.23	4.41	5.4	32.9	2.5	7.9	$1.6\cdot10^{-5}$	$5\cdot10^{-3}$
Cement+Fe+lim+1% Pyrex	0.45	2.26	3.60	5.0	23.3	2.3	7.7	$6.4\cdot10^{-6}$	$4\cdot10^{-3}$

ness D_{th} increased. Water becomes increasingly transparent for the higher energy neutrons; so the addition of a water layer gives essentially the absorption coefficient for the high energy neutrons in the spectrum. For the cement+limonite+iron absorber (Fig. 9B) a value of $D_{th}=5.4$ cm was obtained without the water shield, but with 15 cm of water in *front* of the block $D_{th}=7$ cm at $r=20$ cm (Fig. 9A). This effect and the observed variations in the half-value thickness of the attenuation curve *in water* (at $r=10$ cm, $D_{th}=4.0$ cm and at $r=30$ cm, $D_{th}=5.5$ cm) showed that considerable care must be exercised in the interpretation of our data. It became necessary to investigate whether or not the half-value thickness of 5.4 cm in the limonite+iron mixture corresponded to the absorption of neutrons from the same part of the energy-spectrum as that which gave rise to a $D_{th}=5.5$ cm in water. The slope of the absorption curve in the limonite+iron block was investigated up to larger thickness of absorber by placing the cement+magnetite block in front of it. $D_{th}=5.4$ cm was obtained for the first 20 cm in the second absorber. At larger distances the intensity started to increase again; the neutron flux being so small that diffusion of almost thermal neutrons through the sides of the absorber became observable.

Thus, it cannot be excluded that the attenuation curve will become less steep with increasing absorber thickness, because of the slowing down process of a small fraction of very high energy neutrons. Insufficient knowledge about the various possible ways in which fast neutrons can be slowed down, e.g., inelastic scattering $(n, 2n)$ reactions, makes a detailed discussion about the form of an absorption curve impossible. However, the comparison of our different absorption curves is justified as long as can be shown that in each absorption experiment the same number of neutrons is slowed down to resonance energy. Thus, the number of neutrons slowed down to resonance energy gives evidence whether in some cases an appreciable number of fast neutrons is not sufficiently slowed down to E_{res} and thus fail to be detected. In Section III we found that

$$F_E=\int_0^\infty I_E(r)dr$$

is proportional to the total number of neutrons slowed down to the energy E. For all absorbers studied F_{res} amounts to *ca.* 3.10^3 (water: $F_{res}=2.73\times10^3$, for cement +iron+limonite+Pyrex; $F_{res}=3.51\times10^3$). This proves that in all cases about the same number of neutrons is slowed down to the silver *resonance* energy. The essentially constant value for F_{res} shows that capture of neutrons faster than E_{res} is in general negligible. The value F_{th} for thermal neutrons differs in some cases by a factor 3, because slow neutrons are absorbed before being detected.

A comparison with other experiments shows that absorption coefficients obtained with Ra+Be sources and those obtained with neutrons from Be+16-Mev protons are almost the same. By plotting $I\cdot r$ from the

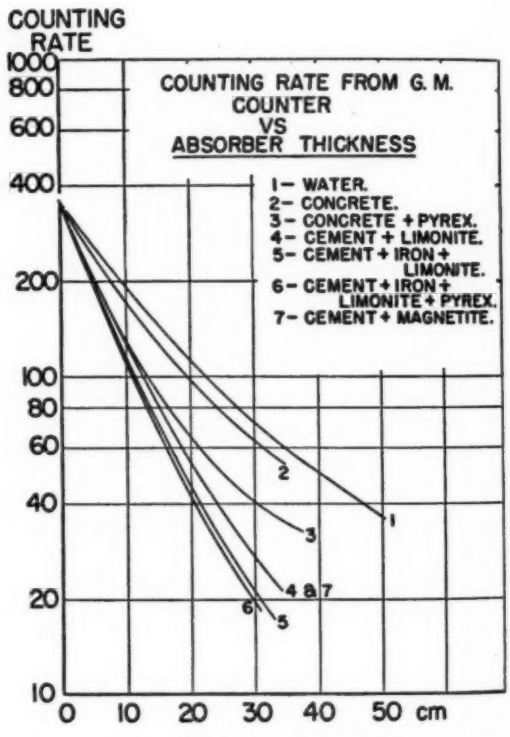


FIG. 11. γ-ray attenuation curves. Abscissa shows the thickness of absorber, ordinate the normalized counting rate of a G-M counter.

TABLE II. Approximate shield thickness for a required attenuation.

		Thickness for neutrons (cm)					Thickness for gamma-rays (cm)				
A[a]	n[b]	Water[c]	Concrete[c]	Magnetite[c]	Limonite[c]	Scrap iron[c] and limonite	Water[c]	Concrete[c]	Magnetite[c]	Limonite[c]	Scrap iron[c] and limonite
10^{-1}	3.33	18.3	35.6	28.3	16.0	17.5	49.4	44.4	30.0	30.0	26.3
10^{-2}	6.66	36.7	71.4	56.6	32.0	35.5	98.8	88.8	60.0	60.0	52.6
10^{-3}	9.99	55.0	107.0	85.0	48.0	53.5	148.2	133.2	90.0	90.0	78.9
10^{-4}	13.32	72.6	142.5	113.2	64.0	71.2	197.6	177.6	120.0	120.0	105.2
10^{-5}	16.65	91.5	178.1	141.5	80.0	89.0	247.0	222.0	150.0	150.0	131.5
10^{-6}	19.98	110.0	214.0	169.8	96.0	107.0	296.4	266.4	180.0	180.0	157.8
10^{-7}	23.31	128.2	250.0	198.1	112.0	125.0	345.8	300.8	210.0	210.0	184.1

[a] A is attenuation factor. [b] n is number of half-intensity values. [c] Columns give thickness in cm to reduce initial intensity by the factor "A."

absorption curve, Rush[18] observed in water with Ra+Be neutrons a half-value thickness of 5 cm is obtained. For plane geometry, Riezler (see Jensen and Ritter[9]) obtained a half-value thickness of 4.9 cm. (For our Be+p neutrons: $D_{th}=5.5$ cm) Jensen and Ritter[9] found a half-value thickness for Ra+Be neutrons slowed down to the In resonance in standard concrete: $D_{\text{In}}=10.6$ cm. (For Be+16-Mev p neutrons: $D_{th}=10.7$ cm.) This is in agreement with the high energy part of the spectra obtained for the neutrons from Ra+Be[12] and for the neutrons from the Be^9+p reaction.

It is apparently difficult to increase the density of an absorber without at the same time decreasing the amount of hydrogen available for slowing down neutrons. Since elastic scattering by a heavy nucleus, such as iron, is very inefficient for slowing neutrons it is of interest to investigate the cross section for inelastic scattering and to compare the resulting mean free path for energy loss with that in water. Barshall, *et al.*[19] found about 10^{-24} cm^2 for the inelastic scattering cross section of iron at 3 Mev, and presumably at higher energies the cross section becomes still larger. At 3 Mev the scattering cross section for hydrogen is $2.3\cdot10^{-24}$ cm^2 and decreases as $1/E$ at higher energies. From these cross sections and the composition of a concrete block containing scrap iron and limonite we find that the mean free path for scattering by hydrogen is 18.5 cm, while the comparable number for inelastic scattering by iron is 28 cm. Pure water has a mean free path of about 6 cm for 3-Mev neutrons but for higher energy neutrons the effectiveness of water falls off; so that above about 5 or 6 Mev the iron in our special concrete should be more effective than pure water. To be most effective against high energy neutrons a shield should therefore include heavy nuclei, but in order to slow down neutrons in the range of about a volt up to 3-Mev hydrogen material must be relied upon. These qualitative conclusions are fully borne out by our experiments.

Table I summarizes the data obtained from our experiments. Column 2 gives the amount of water in g/cm^3 calculated (1), from the water added to the mixture, (2) the moisture content of the components and (3), the water of crystallization of the limonite. Column 3 shows the weight of scrap iron+iron in the limonite per cm^3 of material, column 5 the half-value thickness for the neutron attenuation curves obtained with the Ag thermal neutron detectors. The age of thermal neutrons is calculated in column 6. For all concretes, values for the age are appreciably larger than for water. This is probably due to neutrons of about 1 Mev, at which energy the mean free path for a hydrogen collision is still relatively large and inelastic scattering by the iron improbable. This is especially the case in the cement+magnetite test block.

In Table I columns 9 and 10 give the transmission of a 3-ft. wall calculated for two conditions:

1. For neutrons by the expression:

$$T=\frac{\int_{90}^{\infty} I_{th}(r)dr}{\int_{0}^{\infty} I_{th}(r)dr}$$

(For large thickness of absorber r, and exponential attenuation with half-value thickness D, this may be approximated by $T=2^{r/D}$.)

2. For the G-M counter: The absorption curves for the counter are extrapolated with the absorption coefficient obtained from the slope of the curve at $r=30$ cm.

Column 9 shows that for our neutron spectrum the limonite mixtures are as effective as water in stopping neutrons and that compared to ordinary concrete a 3-foot wall of limonite+scrap iron mixture is 280 times more effective against neutrons. Column 10 shows that against γ-rays the limonite plus scrap iron wall is superior to ordinary concrete by a factor of 20, and in comparison with water there is a 60-fold advantage. It is probable that the superiority of iron+limonite would be even more marked at higher neutron energies.

Although boron is frequently suggested for use in shields on account of strong non-radiative capture of neutrons it is apparent from our results that little is to be gained in this direction with respect to either thermal neutron attenuation or suppression of γ-rays.

The concrete+Pyrex absorber has been tested at 90° to the proton beam. A half-value thickness of 5.9 cm

[18] J. H. Rush, Phys. Rev. **73**, 271 (1948).
[19] Barshall, Battat, Bright, Graves, Jorgenson, and Hanley, Phys. Rev. **72**, 881 (1947).

TABLE III. Measurements on completed shield.

Position in Fig. 12	Shield thickness cm	Ag det. Actual shield	Ag det. Predicted	Counter Actual shield	Counter Predicted
1	0	1[a]	1	1	1
2	30	0.15	0.10	6.8×10^{-2}	5.9×10^{-2}
3	60	$<5\times10^{-2}$	2.4×10^{-2}	3.2×10^{-3}	5.5×10^{-3}
4	90	—	4.9×10^{-3}	5.0×10^{-4}	5.0×10^{-4}

[a] As in the block tests it was necessary to protect the zero absorber silver disk against thermal neutrons coming from the cyclotron by interposing a cadmium filter. The activity of the unshielded detector was 1.77 times larger.

was obtained instead of 7.4 cm in the forward direction, showing that the lower average neutron energy at 90° does indeed lead to larger absorption.

VI. PRACTICAL CONSIDERATIONS

To shield a large installation is always expensive; so practical considerations such as cost, mechanical properties, availability of raw materials and ease with which the shield material can be applied to the actual problem may take precedence over shielding efficacy.

From the mechanical strength standpoint, all the various concrete mixtures showed about the same ultimate compression strength. For example, our tests gave a 28-day strength for ordinary 1-2-3 concrete of 3675 pounds per square inch, while limonite concrete ran 3600 and limonite plus scrap iron about 3000. The early setting properties were normal. Concrete made from scrap iron and limonite is heavy and more difficult to handle in mixing trucks and more difficult to place without voids, but the special problems raised by it are easily solved. The high density of this form of concrete requires sturdy pouring forms and vigorous rodding.

Strength tests showed that the limonite mixtures are adequate for the construction of a radiation shield. An ultimate compression test of cement+iron+limonite cylinders 14 days after casting gave a value of 2193 lb./in.^{2}. For the standard concrete mixture a value of 2210 lb./in.^{2} was observed 7 days after casting.

Limonite ore is widely available, but care must be exercised in specifying high density since some limonite is soft and crumbly. Ore used in our experiments was composed largely of hard, dense lumps ranging from $\frac{1}{8}$ in. to 4 in. to 5 in. A laboratory type crusher reduced the ore to a range of sizes comparable with customary concrete practice. When making up the limonite for incorporation with scrap iron, the ore was crushed somewhat finer than for a pure limonite concrete. A point to notice in specifying limonite is that the amount of absorbed moisture varies considerably with mining conditions and in consequence the amount of water which must be added to make satisfactory concrete will vary somewhat from ordinary concrete practice.

Scrap iron, in the form of small punchings, should be not over $\frac{3}{4}$ in. in size with $\frac{1}{2}$ in. pieces being considerably easier to work. The iron must be grease free and preferably mildly weathered.

Of interest is the comparative cost of the various concretes. Ordinary concrete runs about \$10–\$15 per cubic yard, or about \$10.00 per ton delivered. Limonite ore can be had for \$5.00 per ton at the mine with freight charges increasing the cost to about \$17.00 per ton of ore. Crushing costs are fairly nominal and depend largely on the availability and cooperativeness of local gravel companies. Mixing and delivering of the special concretes, when handled by a commercial concrete firm, may cost as much as the raw materials. Because of the above uncertain factors, it is difficult to quote a reasonable price for limonite concrete, but probably \$40.00 per cubic yard is on the high side. Still more difficult to estimate is the price of scrap iron plus limonite concrete because scrap iron in the most desirable shape is hard to find. The nominal price of scrap iron is presently* \$30–\$40 per ton. Mixing charges are higher for the more dense concretes because of the smaller volume which the mixing trucks can carry. A fair price estimate would be perhaps \$90.00 per cubic yard.

Table II presents the data of Table I in a way which permits cost comparison of different materials for a given attenuation factor. The wall thickness to produce a given attenuation was calculated from the measured half-value thickness on the assumption that there is no change in penetrating power of the radiation for a thickness large compared with 50 cm. Also, for small attenuation the numbers quoted are not accurate because the absorption law is not a simple exponential, e.g., there are *more* thermal neutrons behind 10 cm of concrete than with no concrete at all. With these reservations in mind, one may conclude that to reduce the thermal neutron flux by a factor of one million either water or limonite concrete of one-meter thickness

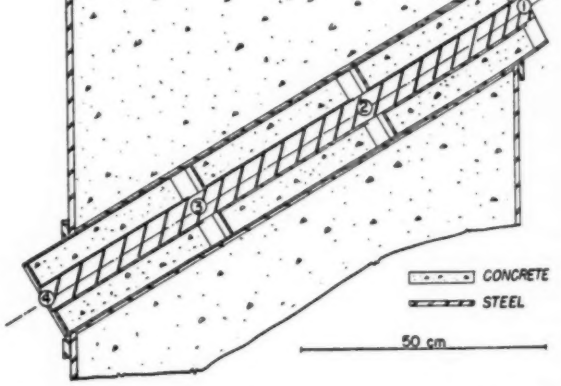


FIG. 12. Cross section through actual cement+limonite+scrap iron shield, showing the port hole, which is directed towards the Be target. The numbers 1 to 4 refer to the positions where the silver detectors and G-M counters were placed to test the actual performance of the shield.

* 1948.

will suffice. Ordinary concrete would have to be about twice as thick. On a cost basis, water is by far the cheapest with straight concrete next and the special concretes much more expensive. On the other hand, a meter of water gives only 10^{-2} attenuation for radiation detectable by a G-M counter while a meter of limonite concrete gives a factor of 10^{-3}, or better, and the scrap iron concrete gives a factor of 10^{-4}. In many installations it is very inconvenient to have a 6-foot wall of water or concrete around the equipment, and in such cases the item of cost is overbalanced by the need for compactness.

Where space and convenience are not at a premium, ordinary concrete is probably the best shielding material. In those regions where limonite, or even magnetite, are readily available and where the transportation, crushing and mixing costs can be kept down it is worth while to use these more costly concretes. On the other hand, the scrap iron concrete is so costly that its use is indicated only where the greater compactness is worth the added price.

VII. PERFORMANCE OF THE COMPLETED SHIELD

Measurements on the finished cement, limonite, and scrap iron shield which has a thickness of 3 feet proved that the neutron and γ-ray attenuation agreed with the extrapolation of the results obtained with the absorber blocks. Figure 12 shows a cross section through that part of the wall in which a port hole has been inserted. The channel is parallel to the neutron flux and in the forward direction of the proton beam incident on the Be target. Three concrete slabs fill the channel. For experimental purposes a 2-in. hole is left in the slabs, which can be closed by a steel rod. For the attenuation measurements, silver disks or a Geiger counter can be placed between the slabs in the usual way.

In Table III the results obtained with the shield are compared with the attenuation expected from the previous experiments. The activity of the Ag detector at position 3 of Fig. 12 was just above the background with the cyclotron off. Only a lower limit for fthe attenuation about 5×10^{-3} could be established at that position; this compares with the value of 2.4×10^{-3} expected from the absorption experiments.

The absorption of the γ-rays appears to have been a little higher than expected, 3.2×10^{-3} against 5.5×10^{-3} at 60 cm. Going from 60 to 90 cm at the outside of the shield the counting rate is only decreased by a factor $\frac{1}{6}$. This is caused by γ-rays transmitted through the doors in the shield which at the time of these measurements are filled with water.

The counting rate for a Victoreen Mica window type counter placed at position 4 of Fig. 12 is 150 counts a minute when a target of Be is bombarded with about 1 μa of 16-Mev protons.

The agreement between observed and empirically predicted attenuation shows that results obtained by our small block tests are sufficiently reliable to permit the design of large shields.

We are indebted to several members of the Princeton faculty for advice on concrete and iron ore problems. In particular, we wish to thank Professor H. F. Winterkorn for his help on concrete problems and Professors A. F. Buddington and E. Sampson for advice on limonite ores. The details of concrete mixing and structural testing were carried out by Professor J. G. Hendrickson of the Engineering School and Mr. B. Setzer of the cyclotron group.

A Bending Method for Measurement of the Tensile Strength and Young's Modulus of Glass Fibers

DAVID SINCLAIR
Johns-Manville Research Center, Manville, New Jersey
(Received August 26, 1949)

The tensile strength and Young's modulus of Johns-Manville experimental drawn glass fibers were measured by a bending method which consists of twisting a loop in a fiber and pulling the ends until the loop breaks. The tensile strength observed by this method was two to three times higher than observed by the conventional straight fiber method, while Young's modulus was practically unchanged. These results may be accounted for by the presence of flaws, since the probability is small that the weakest flaw will occur in the loop, while the straight fiber will always break at the weakest flaw. The apparatus used, a modified Chainomatic balance, is described and the equation of the looped fiber is derived. The method is used to measure the maximum tensile stress and strain attained at the point of minimum radius of curvature of the loop as well as the breaking tensile stress and strain at the actual break point of the fiber. The stress at the break point was found to be as small as one-half the maximum attained in the loop, showing that the strength of the fiber was not uniform.

INTRODUCTION

THE tensile strength of glass fibers is known to vary from a few thousand to two or three million lb./sq. in.[1] depending upon the diameter and length of the fiber, the method of manufacture and amount of handling before test, the age of the fiber and the elapsed time during the test. This wide variation in tensile strength has been supposed to be caused by flaws[1,2] which weaken the fiber at numerous points. The idea is supported by the fact that the tensile strength of fibers increases with decreasing length,[3] since the number of flaws decreases with decreasing length. A very short fiber might therefore be free of any weak flaws and consequently show an unusually high tensile strength.

This paper describes a bending method which provides a measurement of both the breaking stress and strain over a very short length of fiber. The tests were run on Johns-Manville continuous glass fibers made from a plate glass type composition. The method consists of twisting a small loop in a fiber 1 or 2 in. long and pulling the ends until the loop breaks. The tension required to break the loop and the size of the loop at the instant of break are used to calculate the tensile strength at the break point of the loop and the tensile stress attained at the point of minimum radius of curvature of the loop. In addition the extension at these points may be calculated, yielding Young's modulus.

The looped fiber is equivalent to a very short straight fiber since the looped fiber is strained to its breaking point over a length of 1 mm or less. The method is particularly suited to materials such as glass which break under small deformation. In the straight fiber method the extension is too small to be observed conveniently, and short fibers tend to break at the supports. In the looped fiber method the large reduction in the size of the loop during a run magnifies the strain so that it can be readily measured.

[1] C. J. Phillips, *Glass the Miracle Maker* (Pitman Publishing Corporation, New York, 1948), Chapter 4.
[2] F. O. Anderegg, Ind. Eng. Chem. **31**, 290 (1939).
[3] J. J. Bikerman and G. H. Passmore, *The Glass Industry* **29**, 144 (March, 1948).

Murgatroyd[4] used a somewhat similar method to obtain the breaking strain but not the stress of bent fibers. The fiber was bent in the form of a hairpin between the jaws of a miniature vise which was closed until fracture occurred. A maximum extension of about three percent was observed in 25-micron diameter fibers.

METHOD

The apparatus consists of a Chainomatic balance to which is attached a recording pen, and a 22 speed motor-driven Mono-Drum and loading device. Figure 1 shows the recording pen, P, attached to the chain weight drive, W, which is operated by hand in the usual manner. The drum, D, 10 in. long and 6.25 in. in diameter (50 cm circumference), is motor driven at a constant speed of 1 r.p.m., always in the same direction (from right to left under the pen). The looped fiber is hung from the left pan hook of the balance, and its lower end drawn vertically downward at a constant speed of 0.5 in. per min. by the loading rod, R, which is geared to the drum. A gear shift, G, is used to reverse the direct of motion of the loading rod in preparation for a new test. The magnetic damper, M, provides aperiodic damping of the balance.

In order to mount a fiber for test, one end of a straight fiber 1 or 2 in. long is cemented to a small wire hook which is then hung on the left pan hook. A suitable cement is Picein or low melting point De Khotinsky which is melted by means of an electrically heated wire placed against the wire hook. The lower end of the fiber is then cemented to a piece of straight wire held in a pin vise, V, mounted in the loading rod. The pin vise is free to be rotated or raised in the loading rod in order to facilitate formation of the loop. The cemented ends of the fiber are then straightened by applying a slight tension of about 10 mg with the chain weight, and heating each cemented end enough to soften the cement.

With the slight tension still applied to the straight

[4] J. B. Murgatroyd, J. Soc. Glass Tech. **XXVIII**, 368, 388 (December, 1944).

fiber the loading rod is driven up or down (and simultaneously the drum is rotated) until the balance needle is at the zero mark. This point is then located on the drum chart (Fig. 2) by a short vertical line near the top of the chart, drawn by returning the chain weight (with attached pen) to zero. This vertical line locates on the chart the abscissa corresponding to the position of the lower end of the straight fiber. The loading rod is then driven up 0.5 in. corresponding to one revolution of the drum and an abscissa of 50 cm on the chart. A loop is now formed in the fiber by turning the pin vise $1\frac{1}{2}$ to 2 revolutions, raising it until the loop snaps into place and then lowering it to its original position (determined by a shoulder on the pin vise).

The vertical line on the chart now corresponds to the position of the lower end of the looped fiber, which is 0.5 in. above the original position of the lower end of the straight fiber. The equations show (see below) that the elevation, L, of the lower end of the looped fiber above the lower end of the straight fiber is proportional to the minimum radius of curvature, R_m, of the loop, i.e.,

$$L=8R_m. \tag{1}$$

The horizontal line on the chart, drawn during the one revolution of the drum, gives the position of zero force on the fiber. The initially large loop exerts a small downward force, T_0, on the left arm of the balance, which should be compensated by lowering the chain weight until the balance needle is again at the zero mark. The looped fiber is now ready for test.

The test is run by driving the loading rod downward while continuously adding chain weight to keep the balance needle at zero. When the loop breaks, no further weight is added, and the drum is allowed to run slightly to record the breaking tension. The wire carrying the lower portion of the fiber is then removed from the pin vise and saved for measurement of the fiber diameter.*

The tension, T, of the fiber at any point on the curve is proportional to the vertical distance from the point to the horizontal zero line. The constant of proportionality is determined by calibration by weighing. Different ranges were obtained by hanging the chain from different points on the balance arm. The minimum radius of curvature, R_m, is proportional to the horizontal distance, L, from the point to the vertical starting line, measured in the direction of travel of the pen. As described above, one horizontal cm on the chart corresponds to 0.01-in. vertical displacement of the lower end of the fiber.

It is shown below that the maximum tensile stress, σ, in the fiber is:

$$\sigma=\frac{16TR_m}{\pi r^3}=\frac{16TL}{\pi d^3}, \tag{2}$$

and the corresponding extension, e, is:

$$e=r/R_m=4d/L, \tag{3}$$

where r is the radius and d the diameter of the fiber.

Young's modulus of elasticity, E, is:

$$E=\sigma/e=(4TL^2/\pi d^4). \tag{4}$$

The ultimate tensile stress and strain are obtained by measurement from the break point on the chart record.

Some practice is required to keep the balance in equilibrium during a test. The rate of increase of tension is at first very low and then increases until it becomes quite rapid as the break point is approached. The operation is facilitated by observing the balance needle through a magnifying lens (not shown in Fig. 1) of the type manufactured for accurate weighing. The curve obtained is slightly wavy, but the fluctuations are small (usually ± 1 mg), particularly as the loop approaches the break point. Figure 2 shows a representative curve redrawn to reduced scale.

The tension required to break the loop in a 20-micron diameter fiber is of the order of 50 mg. Observations of the looped fiber test performed on a completely automatic, electronically operated Chainomatic balance now available on the market, showed severe hunting of ± 10 mg, resulting in a saw-toothed curve on which the location of a point, in particular the break point, was uncertain. A completely automatic instrument of improved stability has recently been developed in this laboratory and is now being tested.

Equations (2) and (3) give the maximum value of tensile stress and strain, respectively, attained in the loop at the point of minimum radius of curvature, which occurs at the base, O, of the loop (see Fig. 4). Measurements of the break point, as described in the following paragraphs, have shown that the fiber frequently breaks at some other point in the loop where the stress and strain are less than at the base, showing that the

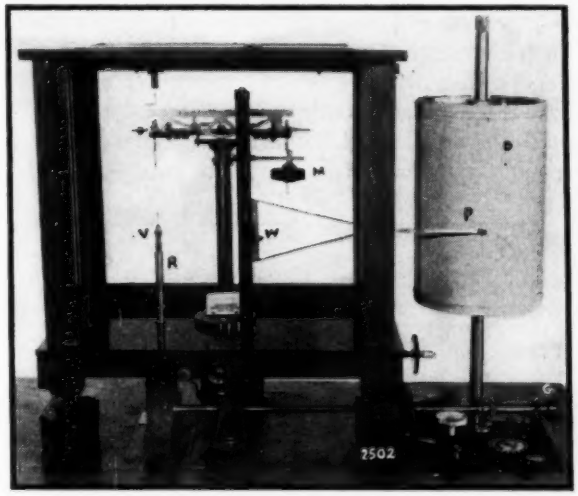


FIG. 1. Semi-automatic fiber balance.

* The diameter was measured with a microscope by placing the broken end of the fiber on a slide in a drop of oil of refractive index 0.05 greater or less than that of the fiber.

strength of the fiber was not uniform. This non-uniformity is not due to variations in the diameter of the fiber. Measurements of the diameter of 20-micron diameter glass fiber samples over a length of several millimeters from the broken end have shown the diameter of any one sample to be constant to two percent. Such a variation in the diameter would cause a variation in the calculated tensile strength of six percent, but the loop has been observed to break at a point where the tensile stress was 50 percent below the maximum attained at the base of the loop.

The measurements of the loop break point were made entirely manually with a preliminary form of apparatus, shown in Fig. 3. The fiber was mounted horizontally with one end fixed in position and tension applied at the other end by a light, circular spring attached to a graduated mechanical slide. The diameter, D, and the position of the loop were observed at the instant of break by means of a low power microscope and micrometer eyepiece. A reading of the mechanical slide was taken corresponding to the straightened fiber with zero tension in the spring, and another at the loop break point. The difference, A, between these two readings is equal to the displacement of the end of the spring attached to the mechanical slide, as shown in Fig. 3.

The displacement, L, is equal to the oppositely directed displacement of the other end of the spring to which the fiber is attached. The equations show that the displacement, L, is proportional to the diameter, D, of the loop (see Fig. 4), i.e.,

$$L=3.75D. \tag{5}$$

Consequently, the total extension of the spring and therefore the tension on the fiber at the instant of break is proportional to $L+A=3.75D+A$. The constant of proportionality was determined by calibration with a Chainomatic balance.

The equations also show that if the fiber breaks at the base of the loop, the remaining portion of the straight fiber (attached to the fixed end) will extend a distance $s=\frac{1}{2}L$ beyond the center of the loop. If the fiber breaks at a point other than the base, the straight fiber will extend a distance $s\pm l$ beyond the center of the loop. Here l is the distance along the fiber from the base of the loop to the break point. Equation (23) shows the tensile stress in the loop as a function of the break point.

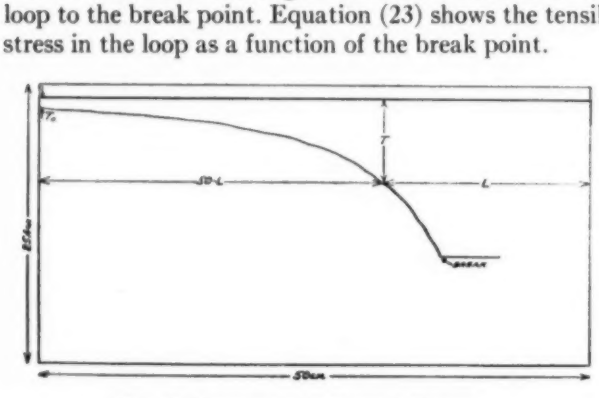


FIG. 2. Chart record of looped fiber test.

After the loop was broken, the fixed portion of the fiber was straightened by wetting the portion attached to the spring and then touching the two portions together along a considerable part of their length. The two portions then clung together so that a slight tension sufficient to straighten them could be applied by the spring. The location of the end of the fixed portion, relative to the recorded position of the loop at the time of break, was read on the micrometer eyepiece. This measurement yields the value of l.

From the above measurements of D and A and the fiber diameter, the tensile stress at the base of the loop can be calculated from Eq. (2). The tensile strength at the actual break point is obtained from Fig. 5, a plot of Eq. (24).

RESULTS

Measurements were made on continuous glass fibers when dry (i.e., not lubricated) and when lubricated with oil applied during the drawing process. Table I gives a summary of the results obtained when using the semi-automatic balance on looped fibers, and a Scott Serigraph on straight fibers. Column 1 shows the type of fiber, column 2 the number of samples tested, column 3 the mean fiber diameter, and column 4 the average deviation from the mean (a.d.). Column 5 shows the arithmetic mean tensile strength, σ_m (at the minimum radius of curvature), column 6 the highest, and column 7 the lowest value of tensile strength and column 8 the average deviation from the mean. Columns 9 to 12 show the corresponding values for the percent extension (also calculated at the minimum radius of curvature), columns 13 to 16 the corresponding values for Young's modulus and column 17 the method.

It is seen that the tensile strength of the dry looped fibers of 20 microns diameter is in average more than twice that of the straight dry fibers. For the lubricated fibers of about 12 microns diameter, the tensile strength of the looped fibers is more than $2\frac{1}{2}$ times that of the straight fibers. No comparison of the extension and Young's modulus could be made between the looped fibers tested on the balance and the straight fibers tested on the Scott Serigraph since the latter apparatus is not sufficiently sensitive to record the small deformation occurring in glass. Previous measurements on straight fibers,† showed the ultimate tensile strain in straight fibers to be one-half to one-third that attained in looped fibers. This ratio is roughly equal to the ratio of tensile strengths obtained by the two methods, so that Young's modulus is about the same for looped and straight fibers.

The straight fibers show an increase in tensile strength with decreasing fiber length. The dry fibers below 1 in. in length are in average 18 percent stronger than those over 1 in. long, and the shorter lubricated fibers are 34

† The straight fiber was broken by applying tension with a calibrated coil spring, and the extension observed with a microscope.

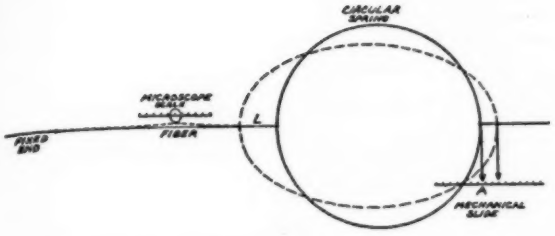


FIG. 3. Schematic diagram of loop break point test.

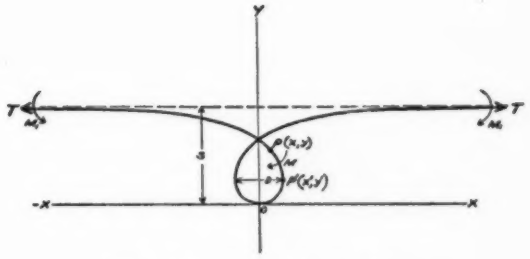


FIG. 4. Diagram of looped fiber.

percent stronger. Straight fibers about 1 mm long would presumably have about the same tensile strength as the looped fibers, provided such short straight fibers could be mounted free of injury or initial strain. This was found to be impossible in practice, the losses of fibers even $\frac{1}{4}$ in. long being excessive. The fiber frequently breaks at one of the cemented ends, or the shock at break causes both portions to fly off since the breaking force of a straight fiber is about 1000 times that of a looped fiber. The losses in the straight fiber test were high even for fibers 1 or 2 in. long. The 39 successful tests of straight fibers over 1 in. long were obtained from a total of 65 tests, while only 38 successful tests of fibers less than 1 in. long could be salvaged from a total of 110. The losses in the looped fiber method were negligible.

No trend in tensile strength with fiber diameter could be observed, because the variation due to flaws was so large. When the tensile strength of the looped dry fibers was plotted against fiber diameter (which ranged from 16 to 32 microns) a typical shot-gun pattern was obtained. The same result was obtained with the lubricated fibers whose diameter ranged from 9 to 16 microns.

Observations of the break point of the loop were made on nine fibers measured with the manually operated spring (Fig. 3). The results, expressed as the ratio of the tensile strength at the break point to the tensile stress at the base, were as follows: 0.999, 0.970, 0.999, 0.890, 0.999, 0.515, 0.997, 0.540, 0.988. These results show that six of the fibers broke very near the base (see Fig. 5), while two of them broke where the tensile strength was little more than one-half the maximum tensile stress attained at the base. In both of the latter tests the diameter of the loop was 0.66 mm, the minimum radius of curvature 0.31 mm and the break occurred within the loop at a distance of about 0.7 mm from its base where the radius of curvature was twice the minimum, or 0.62 mm.

These results show that a region of glass fiber about 1 mm in length can have a tensile strength at least twice as great as that of a nearby region. Additional measurements would probably show a greater variation in the strength of neighboring regions of a fiber. They were not made since this type of measurement is very tedious, and they appear to be unnecessary. The results show nearly a fourfold difference in tensile strength of looped, dry fiber samples. It is to be expected that a sufficiently large number of measurements of the loop break point would eventually show an equally large difference in neighboring regions of some one sample.

DISCUSSION

The results show that the tensile strength of glass fibers increases two to three times when their length is reduced about 75 times. Under the conditions of experiment the variation in the other factors mentioned in the introduction was sufficiently small to have little effect. Although the fibers had a two to one range in diameter the mean diameter was so nearly the same in both types of test that the effect of varying diameter was averaged out. The fibers were taken at random from the same batch and all tests were run at a room temperature of about 75°F and a relative humidity of 45 percent.

Since the fibers were two or three years old and had been roughly handled the observed tensile strength is not representative of the type of glass tested. However, the relative values of tensile strength obtained in the looped and straight fiber tests are in agreement with Anderegg's[2] measurements of the variation of tensile strength with fiber length. Anderegg found that a decrease in length of about 32 times resulted in doubling the tensile strength of 13-micron diameter soda-lime glass fibers. The looped dry fibers (equivalent to a straight fiber about 1 mm long) were $2\frac{1}{4}$ times stronger and the looped lubricated fibers three times stronger than the straight fibers of about 3 in. average length.

The reliability of the 74 tests of looped dry fibers was calculated according to ASTM[5] methods, based on a statistical probability of 0.99. The calculations give tensile strength $\sigma = 2.50 \pm 0.12$ lb./sq. in. (standard

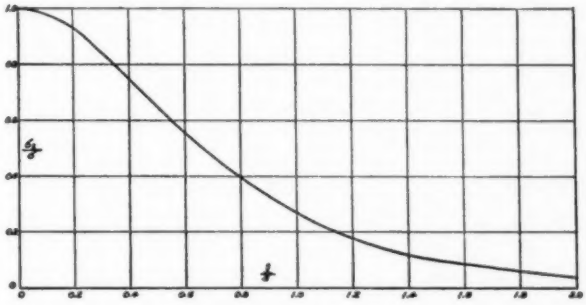


FIG. 5. Tensile stress in looped fiber *vs.* break point. Plot of Eq. (24).

[5] ASTM Manual on Presentation of Data, Supplement A (April, 1945).

TABLE I. Tests of glass fibers.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Fiber type	No. samples tested	Mean fiber diam. microns	a.d.	Tensile strength 10^5 lb./sq. in. Mean	High	Low	a.d.	Percent extension Mean	High	Low	a.d.	Young's modulus 10^7 lb./sq. in. Mean	High	Low	a.d.	Method
								Looped fibers								
Dry	74	20.4	2.4	2.50	3.68	1.55	0.29	3.06	4.76	1.89	0.43	0.83	1.19	0.47	0.10	Balance
Lub.	31	11.3	1.3	3.23	5.29	2.12	0.58	3.40	5.60	1.86	0.68	0.95	2.05	0.50	0.28	Balance
								Straight fibers								
Dry	23	20.3	3.5	1.11	2.01	0.61	0.29	Over 1 in. long								Scott Serigraph
	12	19.5	2.0	1.31	1.64	0.92	0.23	Less than 1 in. long								
Lub.	16	12.2	1.8	1.01	1.48	0.70	0.21	Over 1 in. long								
	26	13.2	1.7	1.35	2.95	0.56	0.47	Less than 1 in. long								

deviation=0.379) and Young's modulus $E=0.825 \pm 0.042$ lb./sq. in. (standard deviation=0.138).

The difference between loading rates in the two tests was relatively small. The loading rate in the straight fiber test was constant at 0.12 lb./min. The rate of stressing, however, varied between 10^5 and 10^6 lb./sq. in./min. depending on the fiber diameter and Young's modulus. The maximum rate of stressing (i.e., in the outermost layer of the fiber) at any time during the looped fiber test can be obtained by differentiating the tensile stress, σ, with respect to the time: [see Eqs. (3) and (4)]

$$\frac{d\sigma}{dt}=E\frac{de}{dt}=-E\frac{e}{L}\frac{dL}{dt}=-\frac{\sigma}{L}\frac{dL}{dt}. \tag{7}$$

dL/dt is negative and constant at 50 chart cm per min. Therefore:

$$d\sigma/dt=50\sigma/L \text{ lb./sq. in./min.} \tag{8}$$

where L is to be measured in cm on the chart as described above. Thus the rate of stressing increases as the test progresses directly with the tensile stress and inversely with the loop diameter.

For example, in Fig. 2 the ultimate tensile stress in this test was 2.70×10^5 lb./sq. in. and the corresponding value of L is seen to be 13.6, from which $d\sigma/dt=10^6$ lb./sq. in./min. At an earlier time during the test when $L=28.7$, $\sigma=1.3\times10^5$ and $d\sigma/dt=2.3\times10^5$. The rate of stressing is of course very small initially but during the latter half of the loop test the order of magnitude of the rate is the same as in the straight fiber test. Anderegg[2] found that a fourfold increase in the rate of stressing of 11-micron diameter soda-lime glass fibers produced no significant increase in the tensile strength. Baker and Preston[6] found that a tenfold increase in the duration of a constant load caused a 15 percent decrease in the bending strength of $\frac{1}{4}$-in. diameter soda-lime glass rods.

As already described, observation of the actual break point of the loop showed non-uniformity of strength over a very short section of a fiber. In most tests the tensile strength was calculated at the minimum radius of curvature, although the loop may have broken at some other point where the ultimate stress was less. In such cases the calculated value of stress was attained at the minimum radius of curvature without breaking at that point, so that this value represents a minimum stress for this point of the fiber. Also, it may be mentioned that the straight fiber losses occurred equally at both high and low loads so that the values reported are a fair average of the tensile strength obtainable in straight fiber tests.

DERIVATION OF EQUATION OF LOOP

Figure 4 shows a diagram of the looped fiber drawn with the base of the loop, the point of minimum radius of curvature, at the origin, O. The loop is maintained by the equal and opposite bending moments, M_1, and tensions, T, exerted at each cemented end by the supports. T acts along a line parallel to the X axis and at a distance, s, from it. In addition there is an equal and opposite torque at each end, exerted about the line through T, to keep the loop from untwisting. As shown below this torque is so small it may be neglected. The looped fiber is symmetrical about the Y axis.

At any point, $P(x, y)$, of the fiber, let the bending moment be M. The portion of the fiber between P and the left cemented end is maintained in equilibrium by the clockwise moment, M, exerted by the remainder of the fiber, and the counterclockwise moments, M_1, and $T(s-y)$.

Therefore,

$$M=M_1+T(s-y). \tag{9}$$

The bending moment $M=EI/R$, where R is the radius of curvature of the fiber at P, E is Young's modulus of elasticity and I is the moment of inertia of the cross section of the fiber.

When the fiber is long compared to the diameter of the loop the bending moment, M_1, at the end of the fiber is negligibly small (see below). Therefore, writing R in

[6] T. C. Baker and F. W. Preston, J. App. Phys. **17**, 170 (1946).

terms of the differential coefficients, Eq. (9) becomes:

$$\frac{1}{R}=\pm\frac{p\,dp/dy}{(1+p^2)^{\frac{1}{2}}}=b(s-y), \tag{10}$$

where $p=dy/dx$ and

$$b=T/EI. \tag{11}$$

The positive sign on the left side of Eq. (10) applies from 0 to P' (at which point $dy/dx=\infty$), where R is positive; and the negative sign applies from P' to the end of the fiber, where R is negative.

Integrating Eq. (10) gives:

$$\pm 1/(1+p^2)^{\frac{1}{2}}=\tfrac{1}{2}by^2-bsy+c. \tag{12}$$

At the origin $y=0$ and $dy/dx=p=0$. Therefore, using the positive sign, $c=1$.

At the cemented end $y=s$ and $dy/dx=0$. Therefore, using the negative sign:

$$b=4/s^2 \tag{13}$$

so that Eq. (12) becomes:

$$\pm 1/(1+p^2)^{\frac{1}{2}}=(2y^2/s^2)-(4y/s)+1. \tag{14}$$

Integrating Eq. (14) gives:

$$x=\pm s\left(\frac{2y}{s}-\frac{y^2}{s^2}\right)^{\frac{1}{2}}\mp\frac{s}{4}\log\frac{1+(2y/s-y^2/s^2)^{\frac{1}{2}}}{1-(2y/s-y^2/s^2)^{\frac{1}{2}}}. \tag{15}$$

The upper signs apply to the left half of the curve and the lower signs to the right half. Equation (15) has been plotted and its form found to fit closely to the observed form of the loop as shown below.

The measurement of the tensile strength by the method described, requires the length of the fiber, l, which is found to be:

$$l=-\frac{s}{4}\log\frac{1+(2y/s-y^2/s^2)^{\frac{1}{2}}}{1-(2y/s-y^2/s^2)^{\frac{1}{2}}}. \tag{16}$$

Equation (16) gives the length of the fiber, measured from the origin along the fiber to a point having the ordinate y.

Equation (2) gives the tensile strength in terms of L, the total length of the fiber minus the distance between the ends of the looped fiber. The length of the fiber between the points (x, y) and $(-x, y)$ is $2l$, given by Eq. (16). The linear distance between these two points is $2x$, given by Eq. (15). Subtracting Eq. (15) from Eq. (16) and multiplying by 2 gives:

$$2l-2x=2s[(2y/s)-(y^2/s^2)]^{\frac{1}{2}} \tag{17}$$

for negative values of x on the left half of the curve and for positive values of x on the right half of the curve.

The difference, $2l-2x$, is equal to the distance L when $y=s$, the ordinate of the end of the fiber. Although, from Eqs. (15) and (16) the abscissa of the end of the fiber and the length of the fiber are both infinite, their difference is finite as shown by Eq. (17). In any practical case the equations apply when the length of the fiber is more than 10 times the diameter of the loop.

From Eq. (17) when $y=s$,

$$2l-2x=2s=L. \tag{18}$$

Equation (4) for Young's modulus of elasticity can now be obtained from Eqs. (11), (13), and (18), noting that for a fiber of circular cross section of diameter d, the moment of inertia $I=\pi d^4/64$.

In the outermost layer of the fiber at a point where the radius of curvature of the loop is R, the extension $e=r/R$.

The minimum radius of curvature, R_m, which occurs at the base of the loop, is obtained from Eq. (10) when $y=0$, i.e.,

$$R_m=1/bs=s/4=L/8 \tag{19}$$

so that the maximum tensile strain in the fiber is:

$$e=r/R_m=8r/L=4d/L \tag{20}$$

corresponding to Eq. (3).

Equation (2) for the maximum tensile stress attained in the loop is given by the product of Eqs. (3) and (4).

The break point of the loop is obtained with the aid of Eq. (18) which shows that the length of the looped fiber is $2s$ greater than the distance between its ends. This means that if the looped fiber were broken at the base (the origin), each straightened portion would extend a distance, s, beyond the Y axis.

Suppose the fiber broke at a distance, l, from the origin, where l is given by Eq. (16). Solving Eq. (16) for y gives:

$$y=s\pm 2se^{(2l)/s}/(e^{(4l)/s}+1). \tag{21}$$

From Eq. (10) when $y\neq 0$, the strain at any point other than the base is:

$$e=(2d/s^2)(s-y). \tag{22}$$

Therefore from Eq. (4), (21), and (22) the maximum tensile stress, σ_l, at a point in the loop a distance, l, from the origin is:

$$\sigma_l=\frac{32TL}{\pi d^3}\,\frac{e^{(2l)/s}}{e^{(4l)/s}+1}. \tag{23}$$

From Eqs. (2) and (23) the ratio of the tensile strength at the break point to the tensile stress at the base is:

$$\sigma_l/\sigma=\frac{2e^{(2l)/s}}{e^{(4l)/s}+1}. \tag{24}$$

Figure 5 shows a plot of Eq. (24) for values of l/s from 0 to 2.

In making the measurements of the loop break point the diameter, D, of the loop is measured across the loop between the points, $y=y'$, where $dy/dx=\infty$ (see Fig. 4). Therefore from Eq. (14):

$$y'=s(1-1/\sqrt{2}) \tag{25}$$

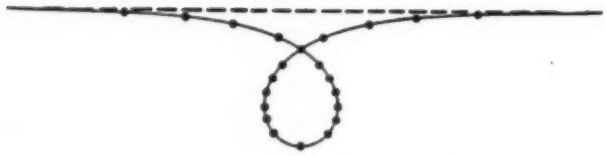


FIG. 6. Photograph of glass fiber loop. Full line—photograph. Circles—calculated points.

so that from Eq. (15), $x'=0.2664s$. Therefore, since $x'=\frac{1}{2}D$

$$D=0.5328s. \tag{26}$$

The assumption that M_1, the moment at the cemented ends, is zero is found to be justified by observation. Since $M=EI/R$, this assumption means that the fiber is straight at the ends. Also from Eq. (15) when $y=s$ this assumption requires a fiber of infinite length. Practically this means that the fiber should be long compared to the size of the loop.

These conditions have been found to be closely maintained in the loops observed. The end portions of the fiber lie along the same straight line to a considerable distance from the cemented ends. (See Fig. 6.) With a fiber 2 cm long the cemented end is located at a distance from the loop of at least 10 times its diameter.

At any point in the fiber, the shearing force, $F=dM/dy=-T$ [from Eq. (9)]. The maximum shear stress occurs at the point P' (Fig. 4) where the shearing force acts perpendicular to the axis of the fiber. At this point the average shear stress, $t=T/\pi r^2$. For a fiber of 10 microns radius and a breaking tension of 50 mg the average shear stress is 225 lb./sq. in. The maximum shear stress, which occurs at the center[7] of the fiber is 310 lb./sq. in. This is negligible compared to the tensile strength of 10^5 lb./sq. in. or more.

The maximum shear in the fiber resulting from the one and one-half to two turns given the fiber to form the loop is produced by one-half to one turn, since one turn is taken out when the loop forms. For a 10-micron radius fiber about 2 cm long the maximum shear strain produced by one turn is 0.3 percent. In most tests one-half residual turn is sufficient to maintain a loop in a fiber 4 or 5 cm long, for which the torsional shear would be less than 0.1 percent.

The one-half turn remaining in a fiber is maintained by equal and opposite couples, C, acting about the cemented ends. The couple, $C=fks$, where f is the normal force between the two halves of the loop at the cross-over point, and ks is the ordinate of the cross-over point. k is a constant $=0.708$. For constant torque in the fiber, C is constant so that the normal force $f=$ const. $\times 1/s$. It was observed that the resulting friction force at the cross-over point was appreciable at the beginning of a test. However, toward the end of the test the friction becomes negligible compared to the tension since the latter increases as $1/s^2$ [Eq. (4)—assuming constant Young's modulus] while the friction force increases as $1/s$. The friction was measured by contracting the loop to near the break and then reversing the direction of the loading rod so as to run the test backward and expand the loop. The friction is added to the tension while the loop is being contracted and subtracted while the loop is being expanded. As a result the return curve tends to lie slightly above the initial curve (of the type shown in Fig. 2), the divergence being significant only at the smaller values of T. In this way the force of friction was found to be about one percent of the tension at or near the break.

[7] S. Timoshenko, *Theory of Elasticity* (McGraw-Hill Book Company, Inc., New York, 1934), p. 290.

EXPERIMENTAL CHECK OF LOOP EQUATION

The ratio, $s/D=1.877$ [Eq. (26)] was measured using glass and Pyrex fibers. The size of the loop was varied by varying the tension over a range of values less than the breaking tension. Using two 20-micron diameter glass fibers, the average of ten measurements gave $s/D=1.887\pm0.005$, 0.5 percent higher than the correct value. Using two 6-micron diameter Pyrex fibers the average of seven measurements gave $s/D=1.870\pm0.015$, 0.4 percent less than the correct value.

A photograph of a 20-micron diameter glass fiber having a loop of 2 mm diameter is shown in Fig. 6. At the edges of the photograph the ends of the fiber are seen to be practically straight and lie on the same straight line (the dashed line drawn to pass through the cemented ends which lie outside the photograph at a distance of one and one-half times the total breadth of the photograph). The ratio $s/D=1.88$ as nearly as can be measured on the photograph. Using the measured value of s, the form of the curve was calculated from Eq. (15). The plotted points lie exactly on the curve as nearly as could be drawn.

ACKNOWLEDGMENT

The writer is greatly indebted to Mr. Paul Greebler of the Physics Section for his assistance in developing the method, and to Mr. H. R. Pittenger of the Physics Section and to Messrs. D. T. Colton, Jr. and Stewart Hayes (formerly of the Physics Section) for running the tests.

Magnetic Field Calculations for Large Cross-Section Cloud-Chamber Coils*

H. W. Koch†
Physics Department, University of Illinois, Urbana, Illinois
(Received September 6, 1949)

A method is described and equations are presented which permit the calculation of the magnetic field characteristics of large cross-section air core coils. The method is applied to the determination of the field non-uniformities on the median plane to be expected with annular cloud-chamber coils of rectangular cross section. The field non-uniformity results are provided in graph form to facilitate the design of cloud-chamber coils.

INTRODUCTION

LARGE, air-core magnetic field coils of the rectangular cross-section type shown in Fig. 1(b) facilitate the study of high energy nuclear particles in the 10 and 100 million electron volt energy ranges. The coils have no windings on the median plane and are used in conjunction with cloud chambers where light and particle beams are admitted on the median plane.

This paper will describe calculations of the magnetic field distributions produced by cloud-chamber coils of large cross-sectional area. Greatest emphasis is placed on the field distribution on the median plane and within the inner radius of the coil. For this region the calculations are presented in a table form. Also, curves are given to allow the determination of the field non-uniformity for almost any rectangular coil configuration with speed and a fair degree of accuracy.

In addition, general equations are derived which permit the evaluation of the absolute magnitude of fields anywhere on the median plane, outside as well as inside the coil bounds. The result is applicable to a wider range of instrument than the cloud chamber: for example, to particle accelerators or to particle analyzers.

To date, the design work on air core coils has been concerned largely with the production of relatively weak, uniform fields for cloud-chamber studies. Since the power requirements were usually small, the field region investigated was chosen close to the axis of the coils. The theoretical methods[1] were based on the assumption that the coil cross section was small compared to the coil radius.

Two general procedures for the small cross-section calculations have been used. In the first method, as employed by Maxwell,[2] the necessary correction due to the cross section was made by the use of Taylor's theorem. The field was expanded to give

$$H=H_0+\frac{1}{24}\left(\xi^2\frac{\partial^2 H_0}{\partial x^2}+\eta^2\frac{\partial^2 H_0}{\partial y^2}\right)+\cdots,$$

where H_0 is the value of the field at the point (x, y) due to a number of coincident circular turns of wire and H is the field at the same point due to the same number of turns of wire spread over a rectangular cross section of axial width ξ and radial depth η.

Gray,[3] Dwight,[4] and Ruark and Peters[5] employ the Maxwell method for calculating fields produced by coils of small cross section. Ference, Shaw, and Stephensen[6] have used Maxwell's method as modified by Lyle.[7]

The second method for handling small cross sections was to integrate the magnetic scalar potential over the area of the coil. Dwight[4] described this method in detail for rectangular cross-section coils. Von Zeipel[8] performed the integration over a cross section which is limited by coaxial cylinders and concentric cones coaxial

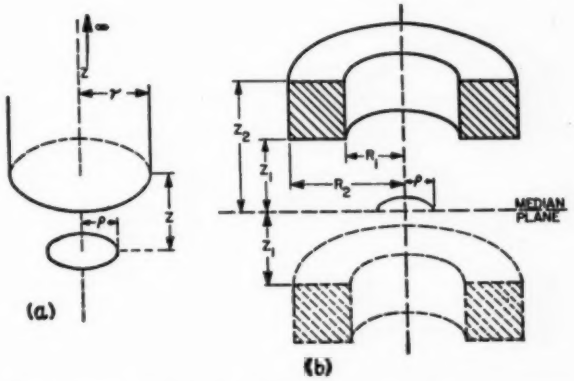


FIG. 1. (a) A semi-infinite solid cylindrical coil of zero core area. (b) The solid rectangular coil above the median plane represents one half of a pair of cloud-chamber coils. It is treated by the addition and substraction of four semi-infinite coils of the type shown in (a).

* This work was assisted by the joint program of the ONR and AEC.

† Present address: Betatron Section, National Bureau of Standards, Washington, D. C.

[1] The calculations to be considered are concerned with circular coils of rectangular copper cross section. The results on rectangular coils of rectangular cross section can be found in the following papers: R. H. Lyddane and A. E. Ruark, Rev. Sci. Inst. **10**, 253 (1939); I. K. Beyerle, Archiv f. Elektrotechnik **25**, 269 (1931); and the experimental paper by E. Nageotte, Comptes Rendus **220**, 557 (1945).

[2] J. C. Maxwell, *Electricity and Magnetism* (The Clarendon Press, Oxford), third edition, (1904), Vol. II, p. 337.

[3] A. Gray, *Absolute Measurements in Electricity and Magnetism*, MacMillan Company, Ltd., London, second edition (1921).

[4] H. B. Dwight, *Electrical Coils and Conductors* (McGraw-Hill Book Company, Inc., New York, 1945).

[5] A. E. Ruark and M. F. Peters, J. Opt. Soc. Am. **13**, 205 (1926).

[6] M. Ference, A. E. Shaw, and R. J. Stephensen, Rev. Sci. Inst. **11**, 57 (1940).

[7] T. R. Lyle, Phil. Mag. **3**, 310 (1902).

[8] E. Von Zeipel, Arkiv. f. Mat. Astr. o. Fys. **30A**, N: 02 (1943).

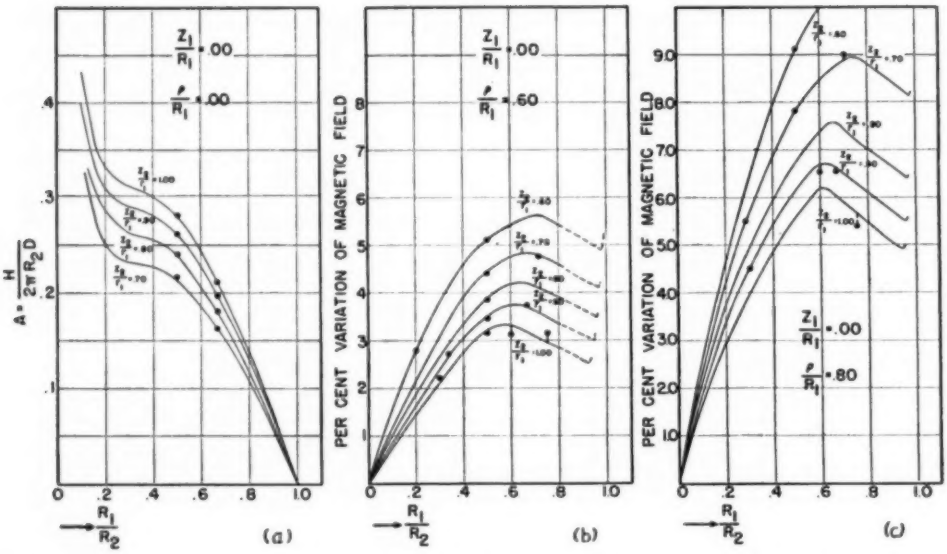


FIG. 2.

FIGS. 2–8. (a) Z component of the magnetic field intensity at $\rho=0$ in gauss-cm per abampere. (b) Z component of the magnetic field variation within a circle of radius ρ for $\rho/R_1=0.6$. The solid points represent calculated points as obtained from Table I. As can be seen, the curve shapes were only partially derived from the points, because of the inaccuracy associated with the points. The specific shapes were obtained from multiple plots and not from individual plots as presented here. The data obtainable from the (b) curves should be accurate to ± 0.1 percent. (c) Z component of the magnetic field variation within a circle of radius ρ for $\rho/R_1=0.8$. The data obtainable from the (c) curves should be accurate to ± 0.2 percent.

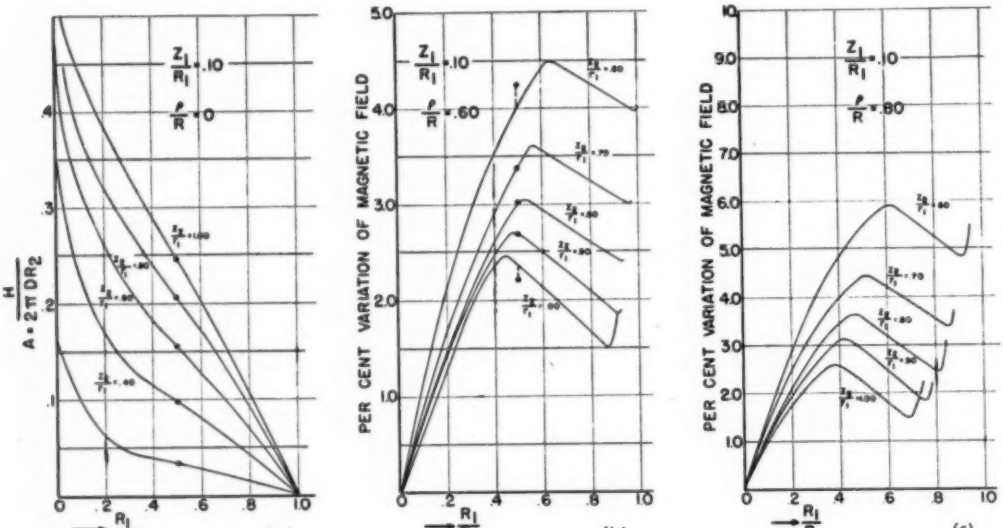


FIG. 3. (See Fig. 2 for caption.)

to the cylinders. Von Zeipel constructed coils to check the calculations.

Foss wrote two papers on the design of air core coils in which the integration procedure is used.[9]

For points close to the air core coils or for points between the coils, the calculations described have limited application because of the slow convergence of the series expressions for the magnetic field or because of the difficulty of integrating exactly over a desired cross section.

Also, calculations are not available on the magnetic field distribution of coils for which the ratio of the inner and outer diameters differ considerably from unity and for which the axial coil height becomes comparable to the radius.

This report will describe a different method for calculating magnetic field distributions, which makes it particularly adaptable to large cross-section coils. The method adopted is to find the field distribution due to a semi-infinite solid cylindrical coil of zero core diameter and unit winding density[10] as shown in Fig. 1(a). By the addition and subtraction of the field of four such semi-infinite coils the field distribution due to the finite hollow coil of Fig. 1(b) can be calculated. The results are still obtained from series expressions. However, the con-

[9] M. H. Foss, Carnegie Institute of Technology, ONR Contract, Technical Reports Nos. 2 and 3 (1948).

[10] The number of ampere turns per unit area of channel section will be called D.

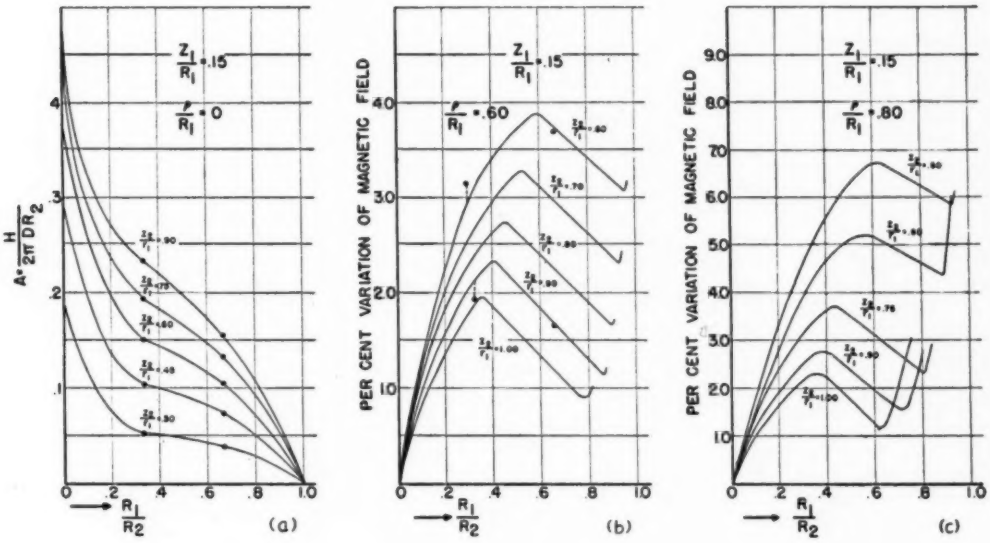


FIG. 4. (See Fig. 2 for caption.)

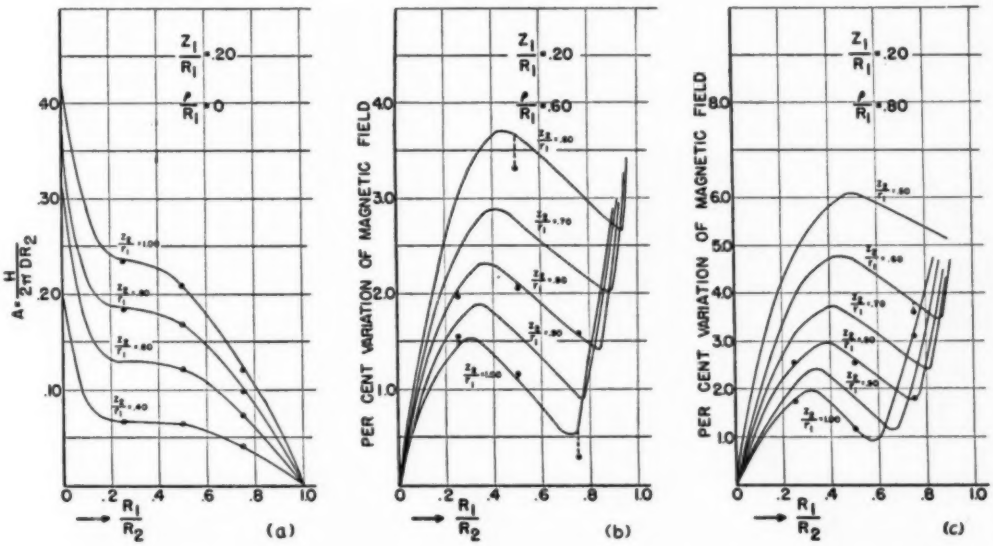


FIG. 5. (See Fig. 2 for caption.)

vergence is reasonable and permits, as an example, the non-uniformity of fields within a measuring circle ρ to be obtained to an accuracy of the order of ± 0.2 percent for ρ's as large as $\rho/R_1 = 0.8$.

The general technique to be described was first used by Butterworth[11] for calculating the mutual inductances for large cross-section coils. The procedure to be followed is very similar to that used by Cockroft[12] in the derivation of his equations. However, the present calculations give magnetic field distributions, while Cockroft used Butterworth's method to determine the mutual inductances and the expected mechanical forces in "Kapitza-type" coils.

THEORY

Consider a one turn coil of radius ρ at a distance of z from the end plane of the semi-infinite, solid coil of radius r, as in Fig. 1(a). The magnetic field in the z direction at any point on the circumference of the one turn coil will be

$$H_z = -\partial\Omega/\partial z, \tag{1}$$

where Ω is the magnetic scalar potential of the solid coil, which must satisfy Laplace's equation at all points external to the solid coil. For this case of circular symmetry about the z axis Laplace's equation can be written as

$$\frac{\partial^2\Omega}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial\Omega}{\partial\rho} + \frac{\partial^2\Omega}{\partial z^2} = 0. \tag{2}$$

A solution of Eq. (2) is

$$\Omega = \int_0^\infty A(\lambda)\epsilon^{-\lambda z} J_0(\lambda\rho)\,d\lambda, \tag{3}$$

where $J_0(\lambda\rho)$ is the zero-order Bessel function given by

$$J_0(\lambda\rho) = 1 - \frac{(\lambda\rho)^2}{2^2} + \frac{(\lambda\rho)^4}{2^2 4^2} - \frac{(\lambda\rho)^6}{2^2 4^2 6^2} + \cdots. \tag{4}$$

[11] S. Butterworth, Phil. Mag. **29**, 578 (1915).
[12] J. D. Cockroft, Phil. Trans. Roy. Soc. **227A**, 317 (1928).

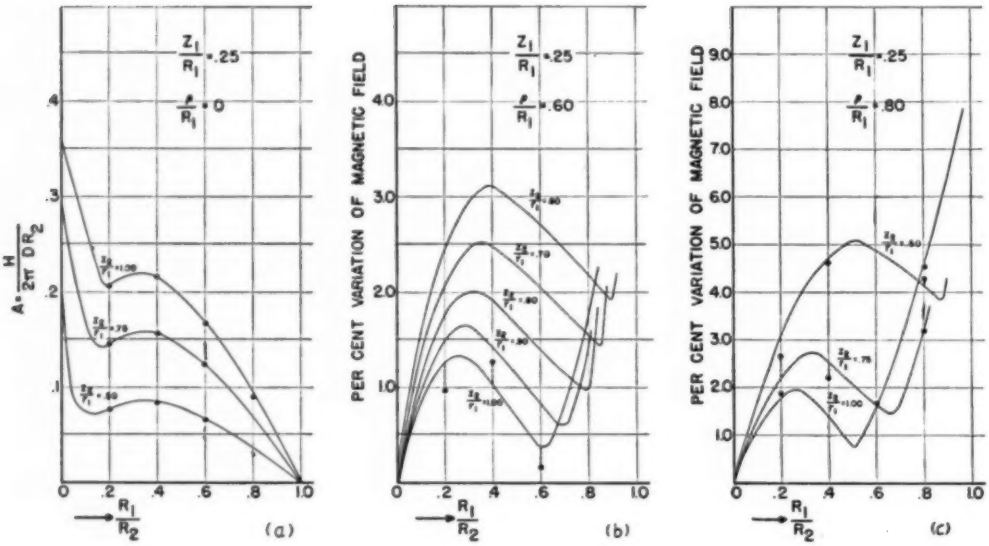


FIG. 6. (See Fig. 2 for caption.)

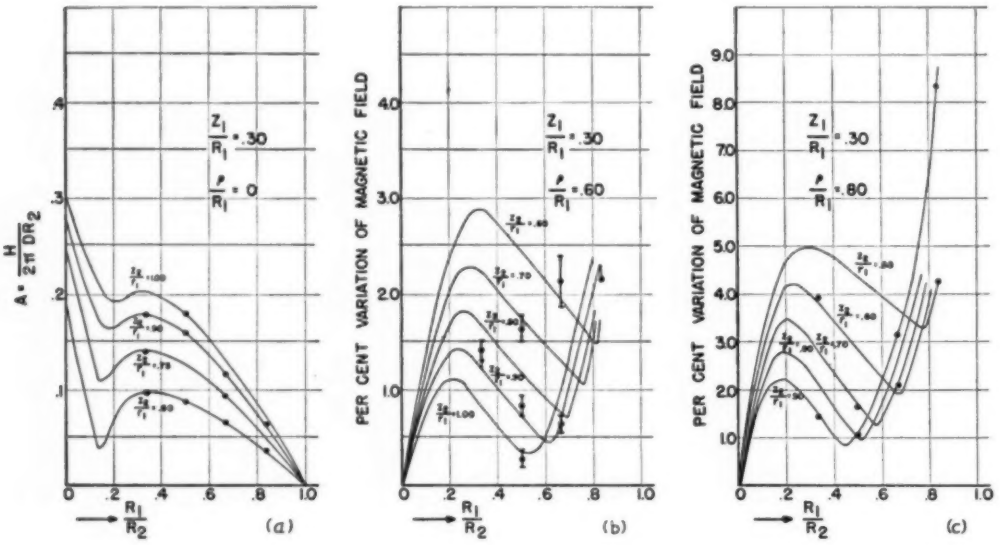


FIG. 7. (See Fig. 2 for caption.)

Case I—For $\rho/r \leq 1$

Since, when $\rho=0$, Eq. (3) reduces to

$$\Omega_0=\int_0^\infty A(\lambda)e^{-\lambda z}d\lambda, \tag{5}$$

Eq. (1) can be written as

$$H_z=\sum_{n=0}^{\infty}(-1)^{n+1}(\rho/2)^{2n}[1/n!]^2\partial^{2n+1}\Omega_0/\partial z^{2n+1}. \tag{6}$$

Ω_0 is the magnetic potential along the axis due to the semi-infinite coil of radius r and D abampere turns per centimeter squared. This is given by

$$\Omega_0=2\pi\int_z^\infty\int_0^r\left[1-\frac{z}{(y^2+z^2)^{\frac{1}{2}}}\right]Ddzdy \tag{7}$$

$$=\pi r^2D\left\{\frac{Z^2}{r^2}\log_e\left(\frac{1+\xi}{z/r}\right)+\xi-\frac{2z}{r}\right\}, \tag{8}$$

where $\xi=[1+(z^2/r^2)]^{\frac{1}{2}}$.

A. For $\rho/r \leq 1$ and $z/r > 1$

The integrand in (7) can be expanded for large z to give:

$$\frac{\Omega_0}{2\pi D}=\sum_{m=0}^{\infty}(-1)^m\frac{(2m+2)!}{[2^{m+1}(m+1)!]^2}\times\frac{1}{(2m+3)(2m+1)}\frac{r^{2m+3}}{z^{2m+1}}. \tag{9}$$

Therefore H_z becomes

$$\frac{H_z}{2\pi D}=\sum_{n=0}^{\infty}(-1)^{n+1}\left(\frac{\rho}{2}\right)^{2n}\left[\frac{1}{n!}\right]^2\times\frac{\partial^{2n+1}}{\partial Z^{2n+1}}\times\left\{\sum_{m=0}^{\infty}(-1)^m\frac{(2m+2)!}{[2^{m+1}(m+1)!]^2}\times\frac{r^{2m+3}/z^{2m+1}}{(2m+3)(2m+1)}\right\}. \tag{10}$$

B. For $\rho/r \leq 1$, $z/r < 1$, and $\rho/z < 1$

H_z from Eq. (6) is

$$H_z = -\frac{\partial \Omega_0}{\partial z} + \frac{\rho^2}{2^2}\frac{\partial^3 \Omega_0}{\partial z^3} - \frac{\rho^4}{2^2 2^4}\frac{\partial^5 \Omega_0}{\partial z^5} + \frac{\rho^6}{2^6 3^2 2^2}\frac{\partial^7 \Omega_0}{\partial z^7} - \frac{\rho^8}{2^8 4^2 3^2 2^2}\frac{\partial^9 \Omega_0}{\partial z^9} + \cdots \quad (11)$$

where the terms $\partial^{2n+1}\Omega_0/\partial z^{2n+1}$ can be determined directly from (8) for small z. It is convenient to separate Ω_0: $\Omega_0 = \Omega_1 + \Omega_2$ where

$$\Omega_1 = \pi D[z^2 \log_e\{r(1+\xi)\} + r^2\xi - 2zr]$$

and (12)

$$\Omega_2 = -\pi D z^2 \log_e z.$$

A substitution of the derivatives of Ω_1 and Ω_2 into (11) results in:

$$\frac{H_z}{2\pi r D} = 1 + \frac{z}{r}\log_e\frac{z}{r} - \frac{z}{r}\log_e(1+\xi) - \left(\frac{\rho}{r}\right)^2 \times \frac{1}{4(z/r)\xi^3} - \left(\frac{\rho}{r}\right)^4 \times 1.5625 \times 10^{-2}\left\{\left(\frac{z}{r}\right)\left(\frac{3}{\xi^5} - \frac{15}{\xi^7} + \frac{1}{\xi^3} - \frac{1}{\xi(1+\xi)^2}\right)\right\} + \left(\frac{\rho}{r}\right)^6 \times 4.3403 \times 10^{-4} \times \left\{\frac{z}{r}\left(\frac{45}{\xi^7} - \frac{735}{\xi^9} + \frac{945}{\xi^{11}} + \frac{9}{\xi^5} + \frac{9}{\xi(1+\xi)^2} - \frac{6}{(1+\xi)^3} - \frac{3}{\xi^3}\right)\right\} - \left(\frac{\rho}{r}\right)^8 \times 2.0345 \times 10^{-5} \times \left\{\frac{z}{r}\left(\frac{525}{\xi^9} - \frac{18585}{\xi^{11}} + \frac{75}{\xi^7} - \frac{45045}{\xi^{15}} + \frac{58905}{\xi^{13}} - \frac{15}{\xi^5} + \frac{15}{\xi^3} - \frac{15}{\xi(1+\xi)^2} - \frac{30}{\xi(1+\xi)^3} - \frac{39}{\xi(1+\xi)^4}\right)\right\} + \left(\frac{\rho}{r}\right)^4 \times 3.1250 \times 10^{-2} \times \frac{1}{(z/r)^3} \times \left\{1 - 3.333 \times 10^{-1}\left(\frac{\rho}{z}\right)^2 + 1.563 \times 10^{-1}\left(\frac{\rho}{z}\right)^4 - 8.75 \times 10^{-2}\left(\frac{\rho}{z}\right)^6 + 5.469 \times 10^{-2}\left(\frac{\rho}{z}\right)^8 - 3.683 \times 10^{-2}\left(\frac{\rho}{z}\right)^{10} + 2.618 \times 10^{-2}\left(\frac{\rho}{z}\right)^{12} - 1.94 \times 10^{-2}\left(\frac{\rho}{z}\right)^{14} \cdots\right\}. \quad (13)$$

C. For $\rho/r \leq 1$, $z/r < 1$, and $\rho/z > 1$

For z less than ρ, the last bracketed term in (13) ceases to converge. This term is derived from the successive derivatives of Ω_2. Thus a distribution of poles must be found which gives the distribution of field

$$H_2 = \sum_{n=0}^{\infty}(-1)^{n+1}\left(\frac{\rho}{2}\right)^{2n}\left[\frac{1}{n!}\right]^2 \times \frac{\partial^{2n+1}}{\partial z^{2n+1}}(-\pi D z^2 \log_e z) \quad (14)$$

in regions for which $z < \rho$. To do this a method similar to that developed by Butterworth is used in which the magnetic field due to an axial distribution of poles extending from $z=0$ to $z=c$ with density $(-\pi z^2 D)$ is found. The field can be shown to be

$$\frac{H_2}{\pi D} = 3z + c - 2(\rho^2+z^2)^{\frac{1}{2}} - 2z\log_e\left[\frac{c+z+(\rho^2+(z+c)^2)^{\frac{1}{2}}}{c+z}\right] + 2z\log_e[z+(\rho^2+z^2)^{\frac{1}{2}}] + \frac{z^2}{(c+z)} - c - \frac{1}{\pi D}F(n, \rho, 1/c). \quad (15)$$

Since H_2 and Ω_2 are independent of the position of c, take $c=\infty$. For $c=\infty$, $F(n, \rho, 1/c)=0$ and

$$\frac{H_2}{\pi D} = 3z - 2(\rho^2+z^2)^{\frac{1}{2}} + 2z\log_e\left[\frac{z+(\rho^2+z^2)^{\frac{1}{2}}}{2}\right]. \quad (16)$$

Equation (16) can now be applied to give H_z for $z < \rho$. Equation (13) for this range of z's is replaced by

$$\frac{H_z}{2\pi r D} = 1 - \frac{z}{r}\log_e(1+\xi) + \frac{z}{r}\log_e\left[\frac{(z/r)+((\rho/r)^2+(z/r)^2)^{\frac{1}{2}}}{2}\right] - \left[\left(\frac{\rho}{r}\right)^2 + \left(\frac{z}{r}\right)^2\right]^{\frac{1}{2}} + \left(\frac{\rho}{r}\right)^2 \times 0.25\left\{\frac{z}{r}\left(\frac{1}{\xi(1+\xi)} + \frac{1}{\xi^3}\right)\right\} - \left(\frac{\rho}{r}\right)^4 \times C + \left(\frac{\rho}{r}\right)^6 \times D - \left(\frac{\rho}{r}\right)^8 \times E, \quad (17)$$

where C, D, and E are terms which correspond to the ones in similar positions in Eq. (13).

TABLE I. $A=H_z/(2\pi Dr)$.

	$\rho/r=0$	$\rho/r=0.1$	$\rho/r=0.2$	$\rho/r=0.3$	$\rho/r=0.4$	$\rho/r=0.5$	$\rho/r=0.6$	$\rho/r=0.7$	$\rho/r=0.8$	$\rho/r=0.9$
$z/r=0.00$	1.00000000	0.90000000	0.80000000	0.70000000	0.60000000	0.50000000	0.40000000	0.30000000	0.20000000	0.10000000
$z/r=0.05$	0.81549483	0.7779395$_3$	0.7071310$_4$	0.6263370$_6$	0.541164$_8$	0.45377$_7$	0.3653$_4$	0.276$_7$	0.188$_7$	0.10$_2$
$z/r=0.10$	0.70017771	0.67795129	0.6262022$_9$	0.5607221$_3$	0.488339$_2$	0.41228$_9$	0.3345$_9$	0.256$_3$	0.179$_7$	0.10$_6$
$z/r=0.15$	0.61062320	0.59531780	0.55600348	0.5025254$_2$	0.4410210$_3$	0.37503$_1$	0.3069$_0$	0.2385$_6$	0.172$_0$	0.11$_0$
$z/r=0.20$	0.53751233	0.52609095	0.49522095	0.4511344$_9$	0.3988146$_7$	0.34166$_9$	0.2822$_6$	0.2225$_7$	0.165$_1$	0.11$_3$
$z/r=0.25$	0.47632186	0.46738471	0.44252487	0.40584188	0.3612455$_0$	0.31181$_2$	0.2600$_1$	0.2081$_0$	0.158$_4$	0.114$_9$
$z/r=0.30$	0.42433107	0.41712368	0.39672451	0.3659468$_4$	0.3278400$_7$	0.28508$_9$	0.2400$_3$	0.1948$_6$	0.151$_9$	0.114$_9$
$z/r=0.35$	0.37970336	0.37377216	0.35679817	0.33079058	0.2981425$_6$	0.26115$_7$	0.22196$_1$	0.1826$_2$	0.1453$_0$	0.112$_8$
$z/r=0.40$	0.34110756	0.33615670	0.32188356	0.29977423	0.2717273$_9$	0.23970$_9$	0.20559$_1$	0.1712$_7$	0.1387$_1$	0.110$_1$
$z/r=0.45$	0.30753199	0.30335642	0.29125683	0.27236551	0.2481706$_4$	0.2204364	0.19072$_1$	0.160$_2$	0.1321$_7$	0.106$_9$
$z/r=0.50$	0.27818229	0.27463289	0.26431000	0.24809757	0.2272327$_9$	0.203114$_2$	0.17718$_9$	0.1509$_1$	0.1257$_6$	0.103$_2$
$z/r=0.55$	0.25241983	0.24933398	0.24053079	0.22656491	0.2084971$_9$	0.187511$_2$	0.16485$_4$	0.1417$_8$	0.1195$_5$	0.099$_3$
$z/r=0.60$	0.22972260	0.22711291	0.21948672	0.20741328	0.1917289$_7$	0.173432$_2$	0.15359$_2$	0.1332$_9$	0.1135$_7$	0.095$_4$
$z/r=0.65$	0.20965837	0.20740547	0.20081111	0.19034086	0.1766910$_3$	0.160704$_8$	0.14329$_6$	0.1253$_9$	0.1078$_7$	0.091$_8$
$z/r=0.70$	0.19186585	0.18991368	0.18419208	0.17508559	0.1631763$_6$	0.149178$_2$	0.13387$_2$	0.1180$_5$	0.1024$_4$	0.087$_7$
$z/r=0.75$	0.17604079	0.17434356	0.16936367	0.16142114	0.1510048$_4$	0.138720$_9$	0.12523$_2$	0.1112$_2$	0.0937$_0$	0.084$_0$
$z/r=0.80$	0.16192561	0.16044551	0.15609850	0.14915252	0.1400197$_4$	0.129214$_0$	0.11730$_4$	0.1048$_7$	0.0924$_4$	0.080$_4$
$z/r=0.85$	0.14930100	0.14800661	0.14420150	0.13811116	0.1300842$_4$	0.120558$_1$	0.11001$_8$	0.0989$_6$	0.0878$_4$	0.077$_0$
$z/r=0.90$	0.13797959	0.13684448	0.13350483	0.12815115	0.1210792$_7$	0.112502$_8$	0.10331$_6$	0.0934$_6$	0.0835$_0$	0.073$_7$
$z/r=0.95$	0.12780081	0.12680262	0.12386391	0.11914614	0.1129007$_6$	0.105438$_2$	0.09713$_9$	0.0883$_4$	0.0794$_0$	0.070$_6$
$z/r=1.00$	0.11862640	0.11774649	0.11515407	0.11098017	0.1054581$_8$	0.098841$_3$	0.09144$_2$	0.0835$_8$	0.0755$_4$	0.067$_6$

Case II—For $\rho/r>1$

Again the methods found in Cockroft's paper can be applied. The flux through a coil of radius ρ is

$$\phi_\rho=\int_0^\rho 2\pi\rho H_z d\rho=M_c I_c, \tag{18}$$

where M_c is the mutual inductance of the semi-infinite coil c in which the current I_c flows. Therefore,

$$H_z=(1/2\pi\rho)(\partial M_c/\partial\rho)\times I_c. \tag{19}$$

However,

$$M_c i_\rho=\int_0^r\int_z^\infty \phi_r D dr dz, \tag{20}$$

where

$$\phi_r=-\int_0^r 2\pi y\frac{\partial\Omega}{\partial z}dy=-\pi r^2\sum_{n=0}^{\infty}(-1)^n\left(\frac{r}{2}\right)^{2n}\times\frac{1}{n!(n+1)!}\frac{\partial^{2n+1}\Omega_0}{\partial z^{2n+1}}. \tag{21}$$

A substitution of (21) in (20) results in

$$M_c i_\rho=\pi D r^3\sum_{n=0}^{\infty}\frac{(-1)^n}{2^{2n}}\frac{r^{2n}}{2n+3}\frac{1}{n!(n+1)!}\frac{\partial^{2n}\Omega_0}{\partial z^{2n}}\bigg|_z. \tag{22}$$

Therefore, from (19),

$$H_z=\frac{Dr^3}{2\rho}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^{2n}}\frac{r^{2n}}{2n+3}\frac{1}{n!(n+1)!}\frac{\partial^{2n}}{\partial z^{2n}}\left[\frac{\partial\Omega_0}{\partial\rho}\right] \tag{23}$$

where Ω_0 is the magnetic potential of a single turn along the axis and is given by

$$\Omega_0=2\pi\left\{1-\frac{z}{(\rho^2+z^2)^{\frac{1}{2}}}\right\}. \tag{24}$$

Since

$$-\frac{1}{2\pi\rho}\frac{\partial\Omega_0}{\partial\rho}=\frac{z}{(\rho^2+z^2)^{\frac{1}{2}}}, \tag{25}$$

H_z is given by

$$H_z=\pi D r^3\sum_{n=0}^{\infty}\frac{(-1)^n}{2^{2n}}\frac{r^{2n}}{2n+3}\frac{1}{n!(n+1)!}\times\frac{\partial^{2n}}{\partial z^{2n}}\left[\frac{z}{(\rho^2+z^2)^{\frac{1}{2}}}\right]. \tag{26}$$

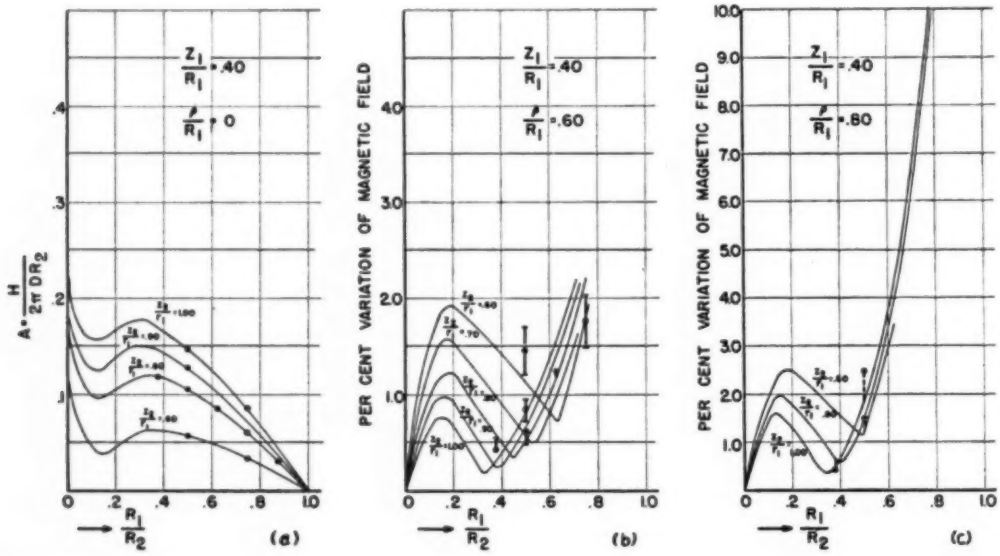


FIG. 8. (See Fig. 2 for caption.)

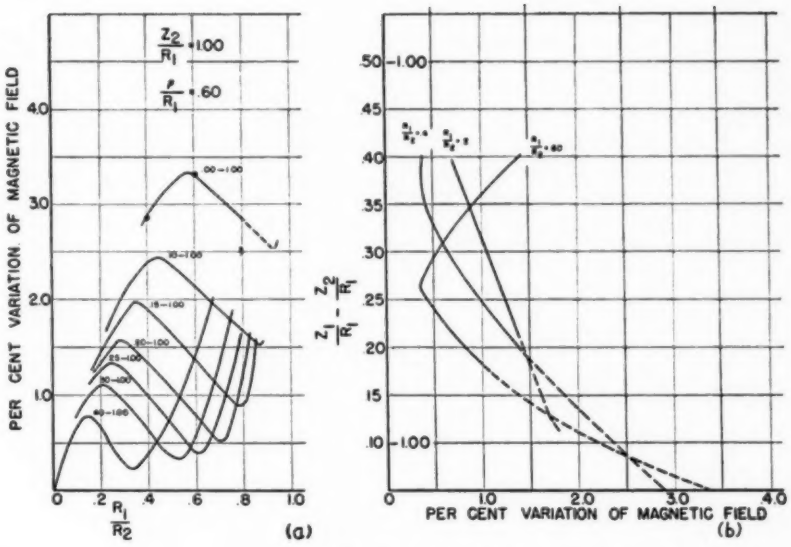


FIG. 9. (a) Percent variation of magnetic field z component variation *versus* R_1/R_2. The constants used for the curves are $\rho/R_1=0.6$ and $Z_2/R_1=1.00$. The variable is Z_1/R_1. The values of Z_1/R_1 and Z_2/R_1 are given above each curve. (b) The data in Fig. 9 (a) have been replotted with the values of Z_1/R_1 and Z_2/R_1 *versus* percent variation of magnetic field.

A. For $\rho/r>1$ and $z/r<1$

Take $\xi=(1+z^2/\rho^2)^{\frac{1}{2}}$. H_z can be evaluated by taking the successive derivatives indicated in (26). The result is

$$\frac{H_z}{\pi D r^3}=\frac{1}{3}\frac{z}{\rho^3\xi^3}-\frac{3r^2z}{20\rho^5}\left[\frac{2}{\xi^5}-\frac{5}{\xi^7}\right]+\frac{r^4z}{336\rho^7}\left[\frac{120}{\xi^7}-\frac{840}{\xi^9}+\frac{945}{\xi^{11}}\right]-\frac{r^6z}{82{,}944\rho^9}\times\left[\frac{5040}{\xi^9}-\frac{68{,}040}{\xi^{11}}+\frac{187{,}110}{\xi^{13}}-\frac{135{,}135}{\xi^{15}}\right]+\frac{r^8z}{8{,}110{,}080\rho^{11}}\left[\frac{72{,}576}{\xi^{11}}-\frac{1{,}596{,}672}{\xi^{13}}+\frac{7{,}783{,}776}{\xi^{15}}-\frac{12{,}972{,}960}{\xi^{17}}+\frac{6{,}891{,}885}{\xi^{19}}\right]\cdots. \quad (27)$$

B. For $\rho/r>1$, $z/r>1$, and $\rho/z<1$

Ω_0 can be expanded in powers of ρ/z. H_z is then obtained by term by term differentiation of Eq. (26).

$$\frac{H_z}{2\pi rD}=\frac{1}{6}\left(\frac{r}{z}\right)^2-\frac{1}{4}\left(\frac{r}{z}\right)^4\left(\frac{\rho}{r}\right)^2\left[1+\frac{3}{10}\left(\frac{r}{\rho}\right)^2\right]+\frac{5}{16}\left(\frac{r}{z}\right)^6\left(\frac{\rho}{r}\right)^4\left[1+\frac{6}{5}\left(\frac{r}{\rho}\right)^2+\frac{1}{7}\left(\frac{r}{\rho}\right)^4\right]-\frac{35}{64}\left(\frac{r}{z}\right)^8\left(\frac{\rho}{r}\right)^6\left[1+2.7\left(\frac{r}{\rho}\right)^2+\frac{9}{7}\left(\frac{r}{\rho}\right)^4+\frac{1}{12}\left(\frac{r}{\rho}\right)^6\right]+\frac{105}{256}\left(\frac{r}{z}\right)^{10}\left(\frac{\rho}{r}\right)^8\left[1+\frac{24}{5}\left(\frac{r}{\rho}\right)^2+\frac{36}{7}\left(\frac{r}{\rho}\right)^4+\frac{4}{3}\left(\frac{r}{\rho}\right)^6+\frac{3}{55}\left(\frac{r}{\rho}\right)^8\right]+\cdots. \quad (28)$$

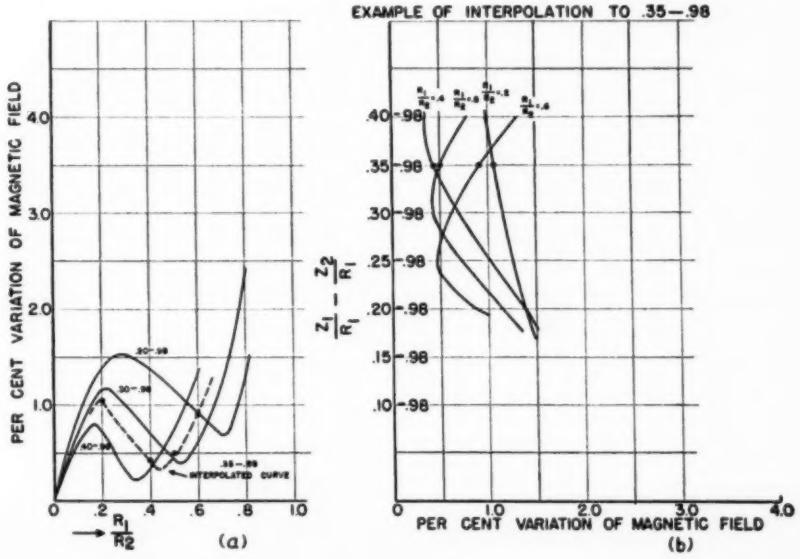


FIG. 10. (a) A plot of percent variation of magnetic field *versus* R_1/R_2 as obtained from the curves in Fig. 10(b). (b) An example of the interpolation to the intermediate values of $Z_1/R_1=0.35$ and $Z_2/R_1=0.98$. The data appearing in this graph were obtained from graphs of Z_1/R_1 and Z_2/R_1 *versus* percent variation of magnetic field [similar to Fig. 10(b)] as obtained from the data in Figs. 2(b), 3(b)···8(b).

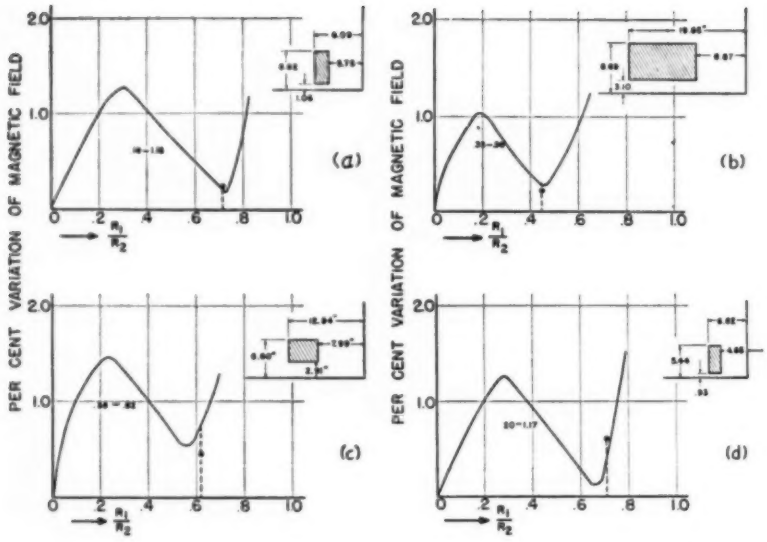


FIG. 11. (a) and (b) The data were obtained from two cloud-chamber coils built at the University of Illinois Betatron Laboratory. (c) The data were obtained from cloud-chamber coils built by Dr. J. R. Richardson in 1939 at the University of Illinois and from the field measurements of Dr. J. A. Phillips, which were privately communicated to the author. (d) The data were obtained from H. R. Crane (see reference 13).

RESULTS

Equations 10, 13, and 17 and Eqs. 27 and 28 permit the calculation of the z component of magnetic field, $H_z(r, z)$, due to a semi-infinite solid coil, whose radius is r and whose distance from the measuring plane is z. Therefore, the field $H_z(R_1, R_2, Z_1, Z_2)$ given by

$$H_z(R_1, R_2, Z_1, Z_2)=H_z(R_2, Z_1)-H_z(R_2, Z_2)-H_z(R_1, Z_1)+H_z(R_1, Z_2)$$

represents the field produced by one finite solid coil of the type shown in Fig. 1(b) and is the result of the addition and subtraction of the fields of four solid semi-infinite coils. The field in the z direction due to two, identical finite coils, one z_1 above the median plane and the other z_1 below the median plane, is twice that due to one coil. There is no radial component of field on the median plane with two coils, because the system is symmetrical.

The five equations enumerated, when the combination principle is used, allow the calculation of the field distributions both inside and outside the bounds of a pair of cloud-chamber coils. Since the present work was primarily concerned with the field distribution close to the axis of cloud-chamber coils and on the median plane, Eqs. 10, 13, and 17 were examined in regions for $\rho/r\leq 0.9$. Table I gives the field values, H_z, divided by $2\pi Dr$ for a semi-infinite solid coil at points for which $0\leq z/r\leq 1.0$ and $0\leq \rho/r\leq 0.9$. The values were all calculated using Eq. 13 and they have limited accuracies as indicated by the number of significant figures listed. Each value is in doubt by less than one in the last figure which in most cases has been subtyped.

The $A(=H_z/2\pi Dr)$ values in Table I were used to examine the absolute magnetic fields at $\rho/R_1=0$, 0.6, and 0.8 for a single finite coil whose dimensions are R_1, R_2, Z_1, and Z_2. In order to provide the results in a usable form, curves were drawn which give the fields on the coil axis [Figs. 2(a), 3(a), $\cdots$8(a)], the field variation within the circle of radius, ρ, for $\rho/R_1=0.6$ [Figs. 2(b), 3(b), $\cdots$8(b)], and the field variation for $\rho/R_1=0.8$ [Figs. 2(c), 3(c), $\cdots$8(c)]. The abscissa scale in each set of curves is R_1/R_2. In the case of the field variations within the circle of radius ρ, the percentage variation was obtained by dividing the difference between the maximum and minimum values of field within the circle of radius ρ by twice the mean value of field. All curves of the sets of three graphs (a, b, and c) apply to a constant Z_1/R_1 and a variable Z_2/R_1 as labeled.

The smooth curves, which are based on the calculated plotted points, should be used for predictions of field distributions. The smooth curves are more accurate than the calculated points, which are obtained from the data of Table I, because all the curves in Figs. 2–8 are internally consistent. This was accomplished by replotting the curves of Figs. 2–8 in inverse curves in which Z_2/R_1 was kept constant instead of Z_1/R_1. Any small inaccuracies in the calculated values were immediately apparent. An example of an inverse plot is shown in Fig. 9(a).

The calculated points of Figs. 7(b) and 8(b) were drawn with their estimated accuracies, in order to show the agreements of the points with the smooth curves.

The inverse plots of Figs. 9(a) and (b) are useful also for interpolation to a combination of z_1/R_1 and z_2/R_1, which is an uneven multiple of 0.05 and does not occur specifically in Figs. 2–8. As an example of interpolation, the field variation was desired for a z_1/R_1 equal to 0.35 and a Z_2/R_1 equal to 0.98. The interpolation was accomplished as shown in Figs. 10(a) and (b), which are inverse plots of the data in Figs. 2–8.

DISCUSSION

Little work has been done with Eqs. 27 and 28 because they deal specifically with points outside the coil bounds ($\rho/r>1$). However, Eqs. 10, 13, and 17, which deal with points where $\rho/r<1$, have been used extensively and several tests have been made to check their reliability. A crude test of Eq. 6, which forms the basis for Eqs. 10,

13, and 17, was made by differentiating with respect to r and z. The result is the formula for a circular current filament, which can be found in any standard book such as Gray's *Absolute Measurements in Electricity and Magnetism*.

A test was made of the values in Table I by calculating the first and second differences in the ρ/r and z/r directions. This served as a test of the numerical computations.

A test of the theoretically predicted values of field found in Table I and plotted in Figs. 2–9 was made by comparing the field variations predicted with those experimentally determined for four cloud-chamber coils.

The result of the test of the curves is shown in Figs. 11(a), (b), (c), and (d).

Figure 11(b) has drawn on it the cross section of one of the pair of coils to which the curve corresponds. As shown, the z_1/R_1 for the coil is 0.35 and the z_2/R_1 is 0.98. The curve in Fig. 11(b) shows the field non-uniformity for $\rho/R_1=0.6$ as a function of R_1/R_2. The procedure for interpolating to the intermediate values of 0.35 and 0.98 was given in Fig. 10(b). The plotted point for $R_1/R_2=0.45$, in Fig. 11(b) was determined experimentally by measuring the relative field values with two flip coils, each driven by $\frac{1}{4}$ H.P. synchronous motors. The experimental value so obtained is good to an accuracy of ± 0.05 percent. Since the theoretical curve should be accurate to ± 0.1 percent, the agreement of theory with experiment is good.

The only other comments to be made about Figs. 11(a), (c), and (d) are concerned with the comparisons of theory with experiment in the latter two figures. In Fig. 11(c) the experimental value is uncertain because of some uncertainty in the coil dimensions. These were obtained by rough measurements on the completed coil.

In Fig. 11(d) the dimensions were obtained from the drawing in the R.S.I. article by Crane[13] which was assumed drawn to scale.

The following procedure is a summary of the methods exemplified above for calculating the magnetic field characteristics for one annular coil whose dimensions are given by Z_1, Z_2, R_1, and R_2:

1. Calculate the ratios R_1/R_2, Z_1/R_1, and Z_2/R_1.

2. The z component of the magnetic field intensity at $\rho=0$ in gauss-cm per abampere (A) is given on the curves of Figs. 2(a), $\cdots$8(a) as a function of R_1/R_2, Z_1/R_1, and Z_2/R_1. For intermediate values which do not appear on the curves, use the interpolation procedure similar to that used in Fig. 10(a) and (b). D is in abamperes per cm^2 and A is equal to H_z for one coil divided by $2\pi DR_2$. The copper space factor of a non-ideal coil is never 100 percent and the correction factor must be included in the evaluation of the true current density, D.

3. The z component of the magnetic field variation within the circle of radius ρ is given as a function of the coil constants in Figs. 2(b) and (c), 3(b) and (c), $\cdots$8(b) and (c). It is believed that the values obtained from the (b) curves ($\rho/R_1=0.6$) are reliable to ± 0.1 percent and the values obtained from the (c) curves ($\rho/R_1=0.8$) are reliable to ± 0.2 percent. The curves should only be applied to coils, for which the ratio of R_1/R_2 is between 0.1 and 0.9.

The tedious numerical computations involved to obtain Table I and the figures in this paper were capably performed by Mrs. Nelson Grisamore and Mr. Norton Moise. The author acknowledges the helpful criticism of Dr. E. F. Shrader, who carefully reviewed this manuscript.

[13] H. R. Crane, Rev. Sci. Inst. **8**, 440 (1937).

The Effect of Compressive Stresses on the Linear Thermal Expansion of Magnesium and Steel

JOSEPH L. ROSENHOLTZ AND DUDLEY T. SMITH
Department of Geology and Mineralogy, Rensselaer Polytechnic Institute, Troy, New York
(Received October 19, 1949)

A series of cylinders of annealed, extruded, pure magnesium and medium soft steel were subjected to a sequence of compressive stresses at room temperature and their respective coefficients of linear thermal expansion in the direction of stress application were determined for the range 20° to 100°C. The cylinders were then heat treated to permit either recovery or recrystallization and the changes in length were measured. Characteristic curves are given for each metal in which the coefficient of linear thermal expansion and length changes are plotted as a function of the true compressive stress. It is shown that these properties may be used to determine the stress history of a test specimen for which characteristic curves are available. It is suggested that this stress analysis technique may find application in the study of rock deformation and in the investigation of metals, both before and after failure.

INTRODUCTION

IN a comparative study of a single crystal of the Iceland spar variety of calcite and Yule marble, in which approximately half of the calcite grains have their optic axes in preferred alignment, it was found[1] that the volume thermal expansion of the marble was twice as great as for the calcite. Furthermore, the calcite showed no change in length after cooling to room temperature from 700°C, while the recovery in the marble resulted in an increase in length in excess of 1 percent. Both phenomena are clearly associated with the deformational stresses of unknown magnitude which produced the marble and which were relieved during the heat treatment. This explanation received substantiation when it was found that a second heating-cooling cycle gave coefficients of expansion of the same order of magnitude as for the unstressed single crystal of calcite.

Before beginning a study of the effects of high, confining, compressive stresses on a series of marble specimens, it seemed desirable to investigate metals in order to ascertain what type of expansion changes, if any, might be expected. Toward this end, extruded, pure magnesium was selected for study because several of its physical and mechanical properties are somewhat similar to those of calcite. Among these are anisotropy, large linear coefficient of thermal expansion, α, and deformation by both twinning and slip. The second metal selected was a medium soft steel containing 0.29 percent carbon and 0.43 percent manganese. While this was nominally a hot rolled steel, metallographic examination gave evidence of preferred grain orientation indicating that the rolling was finished somewhat below the critical point. This was desirable from the point of view of this research because it permitted the study of a two-phase system in which sufficient deformation had occurred to produce some grain distortion in the direction of rolling, which was also the direction of compression and testing.

[1] J. L. Rosenholtz and D. T. Smith, Am. Mineral **34**, 846 (1949).

PROCEDURE

Test specimens of each metal were cut from single bars and were machined to $\frac{3}{8}$ in. diameter and $\frac{3}{4}$ in. long. The magnesium cylinders were annealed for 1 hour at 300°C; the steel cylinders received no heat treatment because it was desired to retain the original structure. The study of a preliminary series of magnesium specimens showed very significant changes which indicated the particular stress regions where more numerous determinations were required. A complete magnesium series was then subjected to progressively greater compressive stresses to the point of rupture at 19 T.S.I. The maximum stress used for the steel corresponded to the region where failure began by extensive plastic deformation in the vicinity of 40 T.S.I.; no fracture occurred up to 135 T.S.I. As a practical expedient, all specimens were allowed to stand at room temperature for at least two weeks to permit recovery to take place.

A specially designed quartz dilatometer having a sensitivity of 2–3 microinches was used for determining α. The equipment, which has been fully described,[1] consists of a transformer bridge and includes a recorder which gives a complete record of dilatation.

Temperatures were measured with a Type K2 Leeds & Northrup potentiometer using a Pt-Pt 10 percent Rh thermocouple. The furnace was precooled to 10°C before measurements were begun and the heating rate was controlled so that the 20° and 100° points were accurately correlated with expansion. In all cases, the values of α are for the range 20° to 100°C.

As will be shown presently, specimens subjected to widely different compressive stresses may have the same value of α. As a result, a unique solution of stress cannot be obtained from an expansion-stress curve alone. Length changes resulting from recovery or recrystallization appeared to give a second useful relationship. The length recovery of magnesium cylinders after heating to 100°C was too erratic to be helpful. However, annealing at 300°C for one hour gave results consistent with the changes in α. Since the steel cylinders showed no change in length after cooling from 100°C, they were

heated in vacuum for one hour at 650°C for relief of stresses. After cooling to room temperature they were again heated in vacuum to 850°C, the length of the cylinders being measured before and after each heat treatment. A consistent set of values was found as a result of recovery but the results of the full anneal were erratic and could not be used.

STEEL

The results of the investigation of the relationship between α and compressive stress of medium soft steel is shown in Fig. 1. True stresses are used since they reveal the actual compression per unit of area for the respective applied stresses.

Smith and Wood[2] determined the stress-strain curve for the atomic lattice of mild steel (0.1 percent C) in compression, their measurements having been made perpendicular to the direction of stress application. Their curve displays some of the same characteristics as Fig. 1 with the principal exception that it reaches a maximum and then remains constant soon after the yield point has been reached. They divided the behavior of the lattice in terms of three phases. In the first phase, in which the compressive stresses are below the yield point, the crystal lattice obeys Hooke's law and recovers completely when the load is removed. The second phase appears at the yield point where an abrupt, permanent, lattice contraction amounting to 0.03 percent was found. In the third phase, which occurs beyond the yield point, they suggest a "regular increase with stress of the permanent residual lattice strain and the superposition of this residual strain on the reversible elastic expansion induced by the applied stress. The two types of strain are so related that their algebraic sum results in an approximately constant lattice spacing at stresses beyond the yield."

Several important characteristics which did not appear in the strain of the lattice of the mild steel referred to above are found in Fig. 1. In the first place, *AB* corresponds with the first phase of Smith and Wood. However, α does not remain constant as it should have if all stresses were relieved when the compressive loading was removed. While the percentage increase in α is small, it nevertheless increases up to the yield point. It is suggested that this effect is produced by the plastic deformation of some of the steel crystals while most of them are reacting elastically. This seems to correspond to the second stage of deformation of polycrystalline aggregates as described by Boas.[3] The residual stresses surrounding those crystals which experienced plastic deformation may therefore account for the increase in α which becomes progressively larger as the yield point is approached.

The second phase is represented by *BC* which is at the yield point. Here an abrupt drop in α was found in excess of 3 percent. This obviously corresponds to the sudden permanent contraction in the lattice. The third phase, corresponding to *CD*, and amounting to only 6 percent of the total stress range of the curve, gave a phenomenally large increase in α and coincides with Smith and Wood's region above the yield point where the lattice expanded.

The most important departure from agreement with permanent lattice changes is found along *DE*, where Smith and Wood obtain a constant value but where α decreases as much as it previously increased. If we assume, as suggested, that a constant lattice spacing is due to the algebraic sum of the permanent lattice expansion and the reversible lattice expansion, then the marked decrease in α must be some measure of the extensive plastic deformation produced by the relatively high compressive stresses. It appears, therefore, that there is a threshold value, represented by the point *D*, where the spacing decreases while the residual stresses become correspondingly greater. This suggestion is consistent with the explanation of the process of recrystallization given by Seitz[4] who explains that the energy stored during cold-work presumably furnishes the driving force for the recrystallization.

There were no changes in the length of the steel cylinders after cooling from 100°C. However, the changes after heating to 650°C for one hour were quite unexpected since, as shown in Fig. 2, there was either no change or a contraction. At the beginning of this investigation, the dimensional changes were observed as a matter of record. However, when it was found that the steel might possess the same value of α for several widely different compressive stresses, a related property was sought for stress correlation whereby a unique solution would be obtainable. This solution, then, is found for the steel in the length changes induced by a stress relief anneal. The inflection in the length-stress curve coincides with the maximum of the expansion-stress curve.

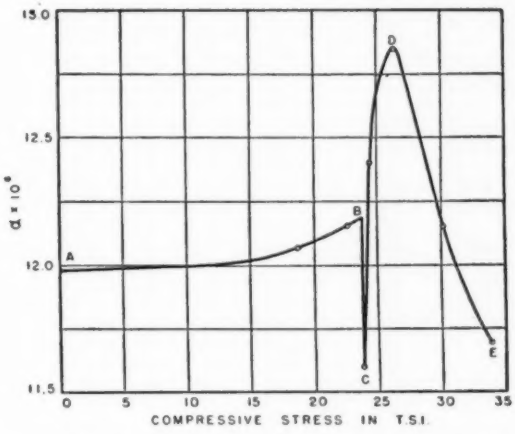


FIG. 1. The relationship between coefficient of linear thermal expansion (20°–100°C) and true compressive stress for medium soft steel.

[2] S. L. Smith and W. A. Wood, Proc. Roy. Soc. **181**, 72 (1942).
[3] W. Boas, *An Introduction to the Physics of Metals and Alloys* (John Wiley and Sons, Inc. New York, 1947), p. 103.
[4] F. Seitz, *The Physics of Metals* (McGraw-Hill Book Company, Inc., New York, 1943), p. 81.

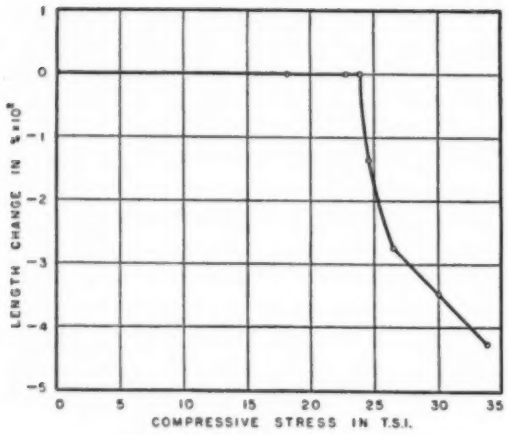


FIG. 2. Percentage change in length of compressed, medium soft steel cylinders after heating for one hour at 650°C.

As for the mechanism of contraction, the abrupt contraction of the lattice spacing at the yield point referred to above coincides with the increase in diameter when plastic deformation begins throughout the steel. As a result, the residual stresses appear to be stored in such a fashion that, when recovery takes place by the heat treatment, expansion of the lattice occurs perpendicular to the direction of compression and thereby causes an actual length decrement.

MAGNESIUM

The effect of compressive stresses on thermal expansion for pure, extruded, annealed magnesium is shown in Fig. 3. The characteristics of the curve are similar to those for steel except for the portion below the yield point. All changes in α at and above the yield point are approximately twice as great as those for steel, but α for magnesium is more than double that of steel. The gradual decrease in α below a true compressive stress of 3 T.S.I. is caused by the fact that measurable permanent deformation occurred, as contrasted with no permanent deformation in steel under a corresponding stress. In other words, appreciable plastic deformation began when the smallest stress was applied and it increased gradually until the abrupt change occurred at the yield point. While it is frequently said that magnesium has no definite yield point, the experimental evidence presented above leads to the opposite conclusion.

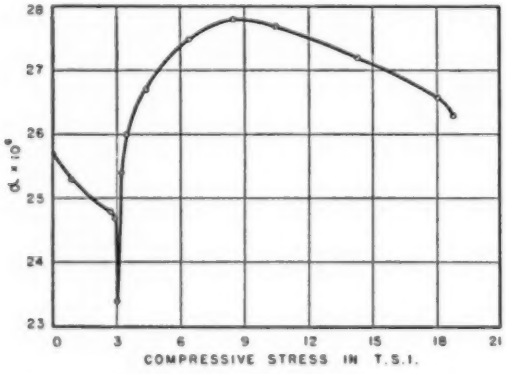


FIG. 3. The relationship between coefficient of linear thermal expansion (20°–100°C) and true compressive stress for pure, extruded, annealed magnesium.

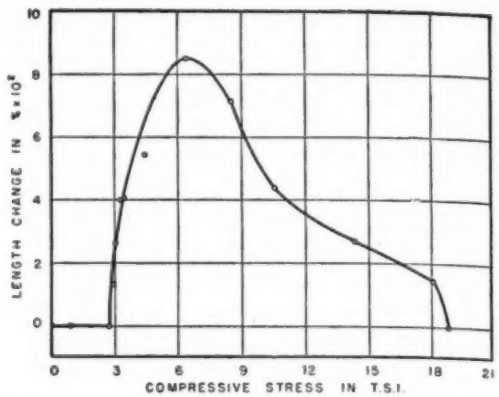


FIG. 4. Percentage change in length of compressed magnesium cylinders after annealing for one hour at 300°C.

As in the case of steel, the same value of α might correspond to two stress values. Measurements after cooling from 100°C could not be used because it was obvious that recovery or recrystallization or both had taken place. Recrystallization may occur in magnesium at 100°C, particularly if it has been stressed. Inasmuch as the different cylinders were stressed to varying degrees, the length changes were too inconsistent to be of use in selecting the correct stress corresponding to a particular value of α.

The results after annealing for one hour at 300°C are shown in Fig. 4. No length changes were measurable up to the yield point, after which the curve is quite similar to Fig. 3. The explanation proposed for the behavior of steel applies equally well to the magnesium. It is possible to arrive at a unique solution of stress from both of these curves except near the peak of the expansion curve where a close approximation would be obtainable.

SUMMARY

It is obvious from the experimental results presented that the coefficient of linear thermal expansion of a material is subject to wide variation and depends upon its previous thermal and deformational history. It is very important to be aware of the fact that if an anisotropic single crystal possesses widely different directional values of α, a polycrystalline specimen of the same material may show an amazingly great variation in α, especially if preferred orientation is pronounced. Disregarding instrumental methods and all attendant factors which pertain to precision of measurement, the published values of α for many materials are so widely divergent as to indicate the importance of knowing their previous history.

The method described above can lead to at least a close approximation of the stresses imposed upon metals

which experience plastic deformation. It is of the utmost importance, however, that the material used in the determination of the characteristic expansion-stress and length-stress curves have the same history as a test specimen for which the unknown superposed stress history is to be determined. This method of stress analysis may be especially useful in the analysis of failure in metals and perhaps in the limit design of structures.

As stated in the introduction, this research was undertaken in an attempt to anticipate how brittle minerals or rocks might react if subjected to compressive stresses under confining pressures. Griggs[5] has already demonstrated that some rocks, such as limestone and marble, behave like metals under high confining pressures. They display an elastic limit and deform by plastic flow when stressed above this point. It is anticipated, therefore, that some minerals and rocks which have been subjected to high, differential pressures will give characteristic curves somewhat similar to those of steel and magnesium. If this proves to be the case, a method will be available for achieving a better understanding of the range of pressures which have been responsible for the great deformational effects produced during metamorphism.

[5] D. Griggs, J. Geol. 44, 541 (1936).

ACKNOWLEDGMENT

The authors are grateful for the assistance provided by the Rensselaer Polytechnic Institute Research Fund. They also express their appreciation for the helpfulness of Professor Joseph F. Throop of the Department of Mechanics and Mr. Stephen E. Wiberley of the Department of Chemistry of Rensselaer Polytechnic Institute.

Thermal Noise at High Frequencies

A. VAN DER ZIEL*

Department of Physics, University of British Columbia, Vancouver, Canada

(Received September 6, 1949)

Spenke's discussion of thermal noise from the point of view of the electron theory leads to a result which differs from Nyquist's original formula at frequencies ν such that $h\nu/kT \gtrsim 1$. This discrepancy is due to the assumptions which Spenke made in his analysis.

1. INTRODUCTION

ABOUT 20 years ago Nyquist[1] published his well-known thermodynamical discussion of thermal noise; since then it has been shown that Nyquist's results could be derived with the help of the electron theory of conduction (Bernamont,[2] Bakker and Heller,[3] and Spenke[4].) Unfortunately Spenke's and Nyquist's results differ at frequencies ν such that $h\nu/kT \gtrsim 1$. It will be shown that the source of this discrepancy is in Spenke's treatment of the problem.

For our discussion we have to define "available power." A signal generator has an e.m.f. e and an internal impedance $Z=R+jX$; its available power P_a is defined as the maximum power which can be dissipated by an external impedance Z'. This maximum occurs if $Z'=R-jX$, so that:

$$P_a=\tfrac{1}{4}e^2/R. \tag{1}$$

2. THERMODYNAMICAL THEORY OF THERMAL NOISE

Nyquist's[1] analysis consists of three steps:

(a) The available noise power P_a in a small frequency interval $\Delta\nu$ of *any* thermal noise generator is a *universal* function of the frequency ν and the absolute temperature T of the generator. He proved this with the help of the second law of thermodynamics.

(b) The available noise power P_a in a frequency interval $\Delta\nu$ around a central frequency ν is:

$$P_a=\bar{E}\Delta\nu, \tag{2}$$

where $\bar{E}$ is the average energy of a harmonic oscillator of frequency ν and absolute temperature T (a system with two degrees of freedom). Nyquist proved this in a somewhat artificial way but his result cannot be doubted. A more detailed proof of (2) was given by Schremp.[5]

(c) $\bar{E}$ is given by the equipartition law. The average energy of an harmonic oscillator of frequency ν and temperature T is:

$$\bar{E}=kT\cdot f(\nu);\quad f(\nu)=h\nu/kT[\exp(h\nu/kT)-1]^{-1}, \tag{3}$$

and one obtains:

$$P_a=kT\Delta\nu\cdot f(\nu). \tag{4}$$

$f(\nu)=1$ if $h\nu/kT \ll 1$; at those frequencies:

$$P_a=kT\Delta\nu. \tag{5}$$

Equation (5) has been verified experimentally; the deviation between (4) and (5) becomes important in the

* Now at: Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota.

[1] H. Nyquist, Phys. Rev. 32, 110 (1928).

[2] J. Bernamont, Ann. d. Physik. Paris 7, 71 (1937).

[3] C. J. Bakker and G. Heller, Physica 6, 262 (1939).

[4] E. Spenke, Wiss. Veroeff. Siemenswerke 18, 54 (1939).

[5] E. J. Schremp, in *Valley and Wallman*, "Vacuum Tube Amplifiers," M.I.T. series (McGraw-Hill Book Company, Inc., New York, 1948), Vol. 18, p. 529.

infra-red (at least at normal room temperature) and no experimental verification of (4) has been published up to now.*

In order to prove (4) in a less artificial way consider a hollow sphere at a uniform temperature T containing an harmonic oscillator of frequency ν_0 and a circuit consisting of a self-induction L, a capacity C and a resistance R and being tuned at the frequency ν_0. The hollow sphere is filled with black-body radiation of temperature T; the harmonic oscillator and the tuned circuit both interact with this radiation and will assume the temperature T. Their average energy E will be the same and the equipartition law will hold. Half of the average energy of the circuit is stored in the condenser, hence:

$$\tfrac{1}{2}C\langle v^2\rangle=\tfrac{1}{2}\bar{E}, \tag{6}$$

where v is the fluctuating voltage across C. A Fourier analysis shows that P_a is proportional to $\Delta\nu$:

$$P_a=\alpha\Delta\nu. \tag{7}$$

We now have to prove: $\alpha=\bar{E}$. Let the thermal noise of the circuit for the frequency interval $\Delta\nu$ be described by an e.m.f. e in series with R, then after (1):

$$\langle e^2\rangle=4P_aR=4\alpha R\Delta\nu. \tag{7a}$$

The contribution of e to the mean square value of the noise voltage v across C in the frequency interval $\Delta\nu$ is:

$$\langle e^2\rangle\left|1+\left(\frac{1}{j\omega L}+j\omega C\right)R\right|^{-2}=4\alpha R[1+4(\omega-\omega_0)^2C^2R^2]^{-1},$$

if $(\omega-\omega_0)\ll\omega_0$ and $\omega_0^2LC=1$. Integrating over all intervals $\Delta\nu$:

$$\langle v^2\rangle=\alpha/C \quad \text{or} \quad \tfrac{1}{2}C\langle v^2\rangle=\tfrac{1}{2}\alpha \quad \text{or} \quad \alpha=\bar{E}.$$

3. THE CORPUSCULAR THEORY OF THERMAL NOISE

The following model of a conductor is used here: Electrons in the conductor move in all possible directions, they have a Fermi velocity distribution at a temperature T. The electrons which have a velocity close to the top of the Fermi distribution can gain energy when an electric field is applied and thus give rise to conduction. Those electrons can collide with the positive ions of the lattice after having traveled an average "free path length" Ω; as they all have practically the same velocity v, the average time of flight τ_0 along the free path length Ω is Ω/v. It is assumed that the collisions with the positive ions of the lattice occur without an exchange of energy and that the electrons are scattered in random directions after such a collision.

The motion of electons along the free path length Ω gives rise to current pulses of length τ_0 in the conductor (thermal noise). Carrying out a Fourier analysis of the resultant noise current $I(t)$, Bakker and Heller[3] and also Spenke[4] found for its Fourier components i:

$$\langle i^2\rangle=\langle i_0^2\rangle(1+\omega^2\tau_0^2)^{-1}, \tag{10}$$

where $\langle i_0^2\rangle$ is the l.f. value of $\langle i^2\rangle$.

If a d.c. field is applied to the conductor, the electrons gain energy in the field; due to the collisions with the lattice a state of equilibrium is established in which the electrons have an average velocity in the direction of the field. This drift velocity, which depends upon τ_0, constitutes a current and so the collisions with the lattice give rise to a finite d.c. conductivity G_0. If an a.c. field is applied, the conductivity G remains independent of frequency as long as its period is small in comparison to τ_0. If it is comparable to τ_0, the electrons will on the average gain less energy in the field, as its phase changes appreciably during the motion along the free path length; at those frequencies $G<G_0$. Spenke[4] found:

$$G=G_0(1+\omega^2\tau_0^2)^{-1}; \tag{11}$$

this result was first obtained by Kronig.[6] Hence:

$$P_a\equiv\langle i^2\rangle/4G\equiv\langle i_0^2\rangle/4G_0. \tag{12}$$

Evaluation of $\langle i_0^2\rangle$ and G_0 then yields (5). This holds for all frequencies.

Spenke's result is not a very surprising one, for in his model no exchange of energy occurs when the electrons collide with the lattice. The average energy stored in the condenser C of an LCR circuit tuned at the frequency ν_0 will then be in equilibrium with the thermal agitation of the electrons. The average energy for translatory motion as given by the equipartition law does not contain the Planck factor $f(\nu)$ and hence one would expect that, independent of the frequency ν_0:

$$\tfrac{1}{2}C\langle v^2\rangle=\tfrac{1}{2}kT$$

so that (5) should be valid for all frequencies. Nowhere in this model is any room left for an expression containing $h\nu/kT$ and this cannot be improved by using another velocity distribution either.

But in a genuine thermodynamical discussion it is not sufficient to state that the electrons have a velocity distribution corresponding to the temperature T; one has to immerse the conductor in a heat reservoir of large heat capacity and one has to investigate how the spontaneous current fluctuations in the conductor are coupled to this heat reservoir. Undoubtedly this coupling occurs as follows: The heat reservoir is in thermal equilibrium with the lattice vibrations and the electrons couple the current fluctuations to those vibrations. This coupling is due to the fact that in each collision an energy exchange $\pm h\nu$ occurs between the colliding electron and the lattice (ν is one of the frequencies of the lattice vibrations). The average energy stored in the condenser C of an LCR circuit tuned at the frequency ν_0

* It would be worthwhile to test its validity in the centimeter wave-length region at low temperatures. If $T=1.2°K$ then $h\nu/kT=1$ at 1 cm wave-length ($\nu=3\times10^{10}$); the deviation between (4) and (5) is then about 40 percent, which might be detectable with present techniques.

[6] R. de L. Kronig, Proc. Roy. Soc. **A124**, 409 (1929); **133**, 255 (1931).

will therefore be in equilibrium with the average energy $\bar{E}$ of the *lattice vibrations* of frequency ν_0, and $\bar{E}$ is equal to the average energy of an harmonic oscillator given by (3). Hence in the correct theory:

$$\tfrac{1}{2}C\langle v^2\rangle=\tfrac{1}{2}\bar{E}$$

so that we obtain (4) instead of (5).

4. THE SOURCE OF THE ERROR

As it is now clear that Spenke's analysis is wrong, it has to be investigated whether the error is in (10) or in (11). According to the Fourier analysis, fluctuating quantities $X(t)$ have Fourier components x_n of frequency ω_n such that:

$$\langle x_n{}^2\rangle=4\Delta\nu\int_0^\infty\langle X(t)X(t+w)\rangle\cos\omega_n w dw. \qquad (13)$$

The correlation function $\langle X(t)X(t+w)\rangle$ is zero if $|w|\ll\tau$, where τ is the *correlation time* of the fluctuating quantity. If $(\omega_n\tau)\ll 1$:

$$\langle x_n{}^2\rangle=4\Delta\nu\int_0^\infty\langle X(t)X(t+w)\rangle dw.$$

This integral is independent of the frequency and is completely determined by the equipartition law; any theory using this law will give the right formula at low frequencies even though it gives wrong values for τ and for $\langle X(t)X(t+w)\rangle$. In the latter case the noise spectrum will have the wrong shape at high frequencies.

Lameris[7] found from his reflection measurements in the infra-red that (11) was correct. He deduced from his measurements that at room temperature $\omega\tau_0$ was equal to unity at the frequencies $\nu(=\omega/2\pi)$: 2.36×10^{13} for Ag, 2.39×10^{13} for Au and 5.23×10^{13} for Pt. But after (4) the available noise power shows a marked decrease if $h\nu/kT=1$, this corresponds to $\nu=0.6\times10^{13}$ at room temperature. Hence the correlation time τ of the fluctuations is about 5–10 times as large as the correlation time τ_0 of the resistance. But in Spenke's model $\tau=\tau_0$ so that the error is definitely in (10).

Spenke[4] made the following assumptions:

(a) After collisions the electrons are scattered in random directions. In fact the scattering over small angles is favored.

(b) No energy is exchanged in the collision process. In fact each collision corresponds to an exchange of energy with the lattice. It is therefore likely that an exact calculation of $\langle i^2\rangle$, in which the above facts are taken into account will give (4) instead of (5). The calculation is very difficult, whereas the thermodynamical analysis is very simple.

In order to show that the two above facts work in the right direction, consider a number of electrons dN having a velocity between v and $v+dv$. In calculating the correlation time of the resistance we have the problem: "What is the probability that the electrons will obtain a different velocity v' when traveling along the distance dx?" This will give a correlation time equal to the time τ_0 between two collisions, if on the average the electrons lose the energy gained by traveling in the electric field after each collision (for a more detailed discussion compare Spenke's[4] paper). In calculating the correlation time τ of the current fluctuations the problem is: "How long does it take before there is no longer any correlation between the current due to the motion of these dN electrons and the original current?" If the scattering over small angles is favored, it will take a number of collisions before this is achieved so that τ is much longer than τ_0, just as was required.

5. AVAILABLE NOISE POWER OF AN ANTENNA

Burgess[8] calculated the available noise power P_a in the frequency interval $\Delta\nu$ of an antenna receiving black body radiation of uniform temperature and showed that it was given by (4). That this leads to (4) and not to (5) can be shown by a simple argument.

At low frequencies the energy density of black body radiation is:

$$U_\nu\equiv(8\pi\nu^2/c^3)\cdot kT\Delta\nu \text{ (Rayleigh-Jeans' law).} \qquad (14)$$

But at those frequencies $P_a=kT\Delta\nu$, hence:

$$U_\nu=(8\pi\nu^2/c^3)\cdot P_a. \qquad (15)$$

Equation (15) describes a property of the antenna which does not depend upon quantum phenomena so that it will hold for all frequencies. But at high frequencies (14) has to be replaced by Planck's law. As this means that (14) has to be multiplied by the Planck factor $f(\nu)$, P_a has to be multiplied by the same factor, so that (4) holds.

The author is indebted to Dr. A. J. Dekker and Dr. F. A. Kaempffer for many discussions on this subject.

[7] P. F. Lameris, Ph.D. Thesis, Groningen, 1936.

[8] R. Burgess, Proc. Phys. Soc. **53**, 293 (1941).

Semiconductive Colloidal Suspensions with Non-Linear Properties

HANS E. HOLLMANN
U. S. Naval Air Missile Test Center, Pt. Mugu, California
(Received October 3, 1949)

Semiconductive particles in a colloidal suspension, when subjected to an electric polarizing field, form semiconductive chains. Once the suspension is polarized, its conductivity increases with increasing fields because the forces of electrostatic attraction press the individual particles together, thus strengthening the transition contacts along the chains. Both effects result in an over-all electric non-linearity that is the subject of extensive investigation.

The non-linear characteristics of such colloidal resistors are oscillographed and described in relation to the polarizing field strength by means of simple formulas based on field coefficients of the first and second order.

In addition to a true rectification, resulting from the curved characteristics, a "pseudo-rectification" occurs which corresponds to an average pre-polarization by stronger alternating fields. The alternating forces of electrostatic attraction assure a high frequency response, so that non-linear Lissajous-figures and harmonics, up to radiofrequencies, are produced.

ACCORDING to Debye's theory of polar molecules the behavior of certain substances, especially of polar liquids and polar dielectrics, such as crystals and ceramics with their abnormally high dielectric constant and permittivity, their rotational and frictional dispersion, their Curie temperatures, etc., is caused by the polarization and orientation of their dipole molecules or their molecule-complexes, when subjected to an electric polarizing field.

For the following considerations it is important to recall that the microcrystalline dipoles originally have a random orientation but, when polarized, orient themselves in the direction of the electric lines of force.

In addition to these polar fluids and substances comprised of natural dipole-molecules, it is possible to enter a more macroscopical region in which the molecules or molecule-complexes are replaced by bigger (but nevertheless microscopically small) particles suspended in a liquid—and later embedded in a solid carrier, preferably in colloidal dispersion. In this way, the concept of polar substances is to be transposed to that of polar suspensions. Attention is drawn to these types of polar liquids by the so-called "magnetic fluid"[1] which represents the prototype in the magnetic field just as ferromagnetism relates to ferroelectricity. For the sake of completeness and easier understanding, it may be convenient to recall the characteristics of the magnetic fluid before we turn to the new polar suspensions in the purely electric field.

The sprinkling of iron filings on a surface under the influence of a magnetic field is a well-known method for making magnetic lines of force visible. All the iron particles become magnetized, attract each other, and arrange themselves in the form of chains in accordance with the magnetic lines of force.

This old phenomenon has been put to ingenious use in the magnetic fluid. Instead of sprinkling onto a surface through free air, the iron particles, or more accurately fine iron powder, is mixed with oil, thus forming an iron-oil suspension, wherein the iron particles are much more mobile than on a solid surface. As soon as the magnetic fluid is subjected to a magnetizing field, it thickens, and if magnetized sufficiently becomes of a paste-like consistency. Hence, the magnetic oil is a new type of a thixotropic fluid with the important difference, that its gel-sol-gel transformation can be controlled by means of an external magnetic field instead of by shearing or stirring.

The solidifying effect can be demonstrated by dipping a permanent magnet into the suspension, after which a certain region around the poles becomes so thickened that it can be pulled out in the form of a semisolid bridge.

The magnetic fluid has found an important application in the magnetic fluid clutch. Instead of utilizing the immediate forces of attraction occurring between the two coupling elements of a mechanical or pure magnetic clutch, a small gap between the coupling elements is filled with the magnetic fluid. Under the influence of a magnetic control field, the magnetic fluid becomes a tough semisolid, binds the coupling elements together, and transmits the torque from the driving to the driven element.[2]

The magnetic fluid, furthermore, may be exploited extensively in electromagnetically controlled hydraulic systems where the controllable thixotropic fluid itself replaces the conventional valves.[3] Another application is for the purpose of forming casting molds where the fluid, in which the model has been immersed, solidifys instantaneously and produces a workable mold as long as it is subjected to a strong magnetic field. Finally, the magnetic fluid may be utilized for forming a controllable resistor similar to a relay. However, such an oil-relay

[1] Jacob Rabinow, "Magnetic fluid clutch," Tech. News Bull. Bur. Stand. **32**, 54–60 (May, 1948).

[2] G. R. Nelson, "Magnetic fluid clutch in servo applications," Electronics 100–103 (November, 1949).

[3] "Magnetic fluid uses," Electronics 120–122 (September, 1949). "New uses for magnetic fluids," J. Frank. Inst. **248**, 155–157 (1949).

will not de-energize by itself but, in order to restore its initial resistance must be stirred or shaken.

Turning to the field of electrostatics, we obtain the so-called "electric fluid."[4] It can be best understood by recalling the so-called "Lichtenberg figures" which in the field of electrostatics are identical to the aforementioned magnetic figures. A light powder having a high dielectric constant, e.g., lycopodium (club moss), in an electrostatic field, forms chains similar to those of the iron filings in the magnetic field, thereby making visible the electric lines of force. Consequently, intermixing or suspending a fine dielectric powder in an insulating liquid of lower dielectric constant—the simplest being in oil—produces the electric fluid whose viscosity also increases transversely to the applied electric field. Hence, it is possible to apply the aforementioned magnetic experiment to the field of electrostatics. Two suitable electrodes or wires, across which either a high direct or alternating voltage is applied, are dipped into the electric fluid. When the wires are withdrawn, a little bridge or membrane adheres. Naturally, as soon as the polarizing voltage is cut off the dielectric suspension assumes its neutral viscosity and drops off.

It is clear that the electrically controlled thixotropy is much weaker than the magnetic chain effect. This is because the electrostatic forces of attraction are limited basically by the breakdown field strength. This disadvantage, however, is compensated for by a very low consumption of control power. Consequently, an electric clutch, of which the torque moment is transferred via the thickened electric fluid, may be of some practical value in cases where a high frequency response and low control power is required.

Fundamentally, the electric fluid and its thixotropy is not completely new. The oldest example is blood. The analogy was observed years ago, in the field of high frequency diathermy.[5] In that part of the body subjected to the diathermy field the blood particles become polarized and also form dielectric chains. Without going into further detail, it may be mentioned that the increased viscosity hampers the blood circulation in the treated vessels, thus counteracting the increased circulation in the vessels which have been enlarged physiologically under the influence of the diathermic heating; hence, the dielectric chains are in constant conflict with the blood circulation, fortunately without endangering the regulative heat flow.

Another characteristic of the electric fluid is the change of its dielectric constant in relation to the applied electric field. A capacitor containing the electric fluid as dielectricum becomes non-linear, i.e., its capacity is a function of the impressed voltage. In this respect, the electric fluid resembles certain polar substances, i.e., rochelle salt, similar crystals, and ceramics with Seignette-electric properties. Referred to the viscosity effect it is obvious that the electric non-linearity occurs long before any mechanical effect can be observed. It would be beyond the scope of this paper to go into greater detail concerning this electric characteristic.

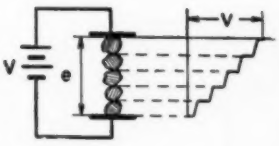


FIG. 1. Qualitative representation of a semiconductive chain and the incremental voltage drop.

After the foregoing introduction, particularly concerning the non-linear capacitor, we proceed to a new concept, namely, an "electric oil" having non-linear or controllable conductivity. For this purpose the dielectric particles must have a certain conductivity which will enable the chain formation to increase the over-all conductivity at the expense of the viscosity effect which now loses its significance. Finally, the electric oil can be mixed or composed in a manner so that its conductivity exceeds its dielectric constant by several orders of magnitude. As a result the wattless component vanishes at the expense of the dissipative component. It is therefore possible to make a non-linear resistor whose value depends on the applied electric field as well as on various external or internal influences. As will be shown in the following pages, the electric oil displays a surprising versatility and creates new fields for numerous applications.

1. THE FORMATION OF SEMICONDUCTIVE CHAINS

In the course of numerous experiments with various mixtures and suspensions, it was found that a suspension of graphite in oil produces very satisfactory results. Besides relatively crude mixtures of coarse graphite powder and lampblack in silicon oil, some colloidal suspensions, such as Oildag, Castordag, and Glydag, operate very stably, particularly so because no sedimentation takes place. Although it is probable that these compositions do not represent an optimum, and taking into consideration that further investigation may disclose better samples, the present investigations are concerned only with graphite-oil suspensions by means of which the unusual behavior and versatility of the electric oil can be proven very satisfactorily.

The process of chain formation can be observed under a microscope. In neutral condition or in the virgin state the particles are distributed irregularly. As soon as the suspension is subjected to a polarizing field, the particles form chains. The picture, however, is not ideal because the primary chains, extending mostly in the direction of the electric field lines, are crisscrossed by secondary

[4] W. M. Winslow, U. S. Patent 2,417,850 (March, 1947). "Induced fibration of suspensions," J. App. Phys. **20**, 1137–1140 (1949).

[5] W. Krasny-Ergen, "Nicht-thermische Wirkungen elektrischer Schwingungen auf Kolloide" (Non-thermal effects of electric oscillations on colloids), Zeits. f. Hochfrequenzt. u. Elektroakustik **48**, 126–133 (1936); Report of the International Congress for UHF-Waves in Physics, Biology, and Medicine, Vienna (1937), pp. 180–184.

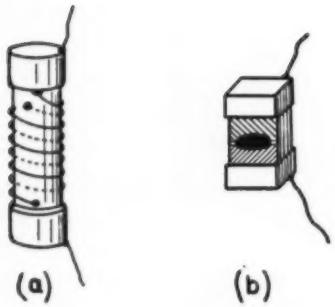


FIG. 2. Two examples of colloidal resistors.

chains extending between points of different potentials along the primary chains, as a result of which a complex network appears instead of individual parallel chains.

The chain formation alone, however, does not suffice to explain the non-linear conductivity. Once the chains are formed at a certain field, a second effect is responsible for the non-linear property, namely, the fluctuating forces of electrostatic attraction along every individual chain, or more accurately, between contacting particles or links. This is explained in greater detail by means of the qualitative representation shown in Fig. 1. The left side illustrates a simple chain between two electrodes with the separation e and the electric field $E=V/e$ where V is the voltage. The right side illustrates roughly the voltage drop along the chain. The semiconductivity of the particles, together with the high transition resistance of the regions where the particles touch each other, effect an incremental voltage drop. Partial fields appear preponderantly at the transition regions, whereas the conductivity prevents higher voltage drops across the particles themselves.

Furthermore, the relatively high fields created in the transition regions produce strong electrostatic forces of attraction between all adjacent particles. As a result, the dielectric chains are reinforced at increasing field strengths so that the increasing over-all conductivity of every individual chain, and therefore of the sum of all chains, may be considered as being caused by the action of innumerable fluctuating contacts between innumerable links. Once the chains are formed, the conductivity continues to increase with increasing field strength, i.e., the over-all conductivity of the electric oil becomes non-linear. A very important result of this contact action is the fact that the non-linearity does not depend on any movement of the particles through the carrier oil but only upon the reinforced pressure along the chains which produces the non-linear conductivity and assures a high frequency response. (A similar effect of strong electrostatic attraction at moderate voltages is known as the Johnson-Rahbek effect[6] occurring in the microscopically small airgap between the surface of a semiconductor, e.g., slate or agate, and an impressed foil electrode.)

[6] K. Rottgardt, "Elektrische Anziehung nach Johnson-Rahbek und ihre Anwendung" (Electrostatic attraction according to Johnson-Rahbek and its application), Zeits. f. Tech. Physik **2**, 315 (1921).

The ratio between the dielectric constant of the particles and that of the carrier oil, the conductivity of the particles, and their size, is also of great importance. It can be seen at once that the foregoing theory leads to an important conclusion. If the individual chain contains, for example, N links, the partial voltage drops between subsequent links can be considered as V/N if the potential drop across the particles themselves, in rough approximation, is neglected. Consequently, noticeable forces of attraction occur only under the assumption that V/N exceeds a certain minimum value. This simple relationship determines the applicability of the electric oil, or, more accurately, its non-linearity and voltage sensitivity. The latter quantities can be chosen by matching the electrode separation to the operating voltages, e.g., by selecting the suitable full operating field strength in relation to the size of the particles.

2. COLLOIDAL RESISTORS

For any practical applications or experimental investigations, the space between suitably arranged electrodes must be filled with electric oil so that a non-linear colloidal resistor is obtained. In analogy to a fluid capacitor, a simple form would consist of two flat or cylindrical electrodes with a fixed or variable separation arranged in an insulated container filled with the oil. If an electrical stability as high as possible is desired, care must be taken that there is no surplus amount of oil outside of the electrodes because the electrid field gradually attracts the dielectric particles from the surroundings so that the mixture ratio in the

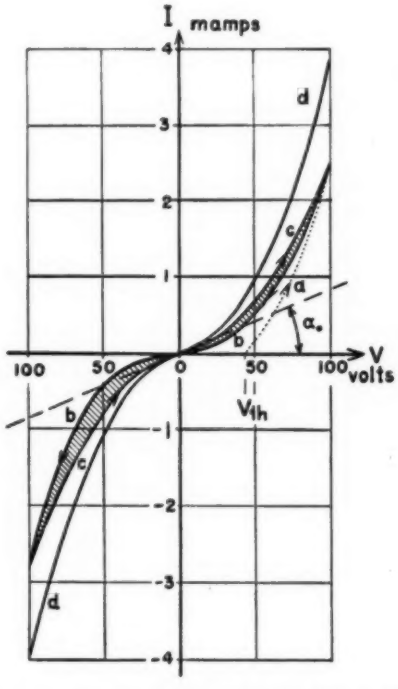


FIG. 3. I–V characteristic of a colloidal resistor at various states of polarization.

inter-electrode space increases at the expense of the surplus oil. Another modification is a glass tube closed at both ends and filled with the oil in which wire electrodes are immersed at each side.

Two more convenient types of oil-resistors are schematically indicated in Fig. 2. Type (a) contains two bare wires which are bifilarly wound around an insulating rod. The non-linearity or voltage sensitivity depends on the separation of the wires, in other words, on the pitch, whereas the resistance range or more accurately, the zero resistance R_0 at very low voltages is related to the surface of the electrodes, i.e., to the length of the bifilar helix.

When a non-linearity as high as possible is required, the second type (b) is very convenient. A minute electrode gap is made by cutting a slit into the silver layer of a little mirror, thus assuring a high polarizing field even at moderate voltages. Naturally, this mirror type may be combined with the former form by cutting a helical slit into a metal-plated rod or by cutting a meandering line similar to a Rowland grating.

Instead of being immersed in an oil-filled container, the electrode-gaps may be covered with a thin layer of electric oil which may be protected by wrapping or hermetically sealing. Whether one type or the other is preferred depends on the various conditions of operation. It must be clearly understood that every colloidal resistor with a larger volume of free oil is more or less microphonic, whereas the sensitivity to vibration vanishes in all cases where only a thin layer of the oil is supported by an insulator.

The possibility of matching the power dissipation as well as the voltage sensitivity to any desired range is of great practical significance. Whereas the power dissipation is determined largely by the cross section or by the current density—if necessary with air or water cooling —the voltage sensitivity or non-linearity, respectively, depends for a given suspension only on the oil gap. Accordingly, the two electrode arrangements shown in Fig. 2 allow encompassing a wide range of non-linearity as well as resistivity.

In addition to the preceding considerations, the colloidal resistor may be compared to the previously mentioned magnetically controlled resistor because both devices have in common the fluctuating transition resistances between individual particles. An important difference, however, is that the electric oil leads to a steadily increasing transition or controllability whereas the magnetic fluid relay operates, so to speak, only in "black-and-white," i.e., its resistance fluctuates only between infinity and zero. Furthermore, the electric oil requires no artificial reduction to its initial state, in contrast to the fluid relay. Only when the electric oil is overloaded do its particles burn and fuse together, and then they must be broken up by stirring, or by strong vibrations.

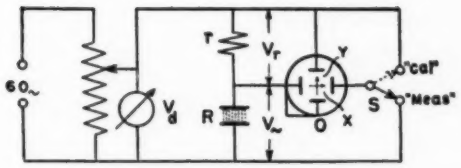


FIG. 4. Simple device for oscillographing non-linear I–V characteristics.

3. NON-LINEAR CHARACTERISTICS

(a) Static Measurements

For obtaining a preliminary insight into the non-linear properties of colloidal resistors, numerous types were investigated at static conditions. In these experiments, the resistors were merely fed by a variable direct voltage V and the passing currents I were observed. Some typical examples of simple I–V characteristics are shown in Fig. 3 in the form:

$$I=f(V)=V/R=G\cdot V, \tag{1a}$$

where R and G are the non-linear resistance and the non-linear conductance, respectively.

A detailed study of the characteristics yields the following results: At the beginning of the measurements, the electric oil is in its virgin state. Because no chains are present, current cannot flow. As soon as the impressed voltage reaches a certain threshold value V_{th}, the chains become active and produce a semiconductive connection between the electrodes. From then on the current rises rapidly with the voltage and follows the steep virgin curve (a) thereby revealing the progressive strengthening of the chains by the increasing forces of electrostatic attraction. When the chains, under the influence of a certain polarization, have attained a certain state, they keep this assumed state, at least to a certain degree, even though the voltage is decreased. In this particular example not enough time has elapsed at the extreme maximum voltage to permit the chains to become fully developed. Consequently, the first I–V characteristics (b) and (c) do not coincide but circumscribe the shaded area.

In order to describe the static picture as completely as possible, it must be emphasized that only a relatively small range can be measured. As noted before, the electric oil becomes unstable at higher current densities because the graphite particles burn and fuse together so that the original colloidal character is completely destroyed.

Quite obviously, the phenomenon of the chain formation has a certain analogy to hysteresis, with the important difference being that the chain formation has a certain inertia. As a result, the chain hysteresis disappears, even at low frequencies, because the chains assume an average state of polarization.

In regard to this time lag and, therefore, to the previous history, the static step-by-step method is somewhat delicate and laborious. A steady state is

reached only after a sufficient time has elapsed before the next step. Accordingly, only after a sufficiently strong polarization has taken place, preferably connection with some reversals of the polarizing field, the fairly stable I–V characteristic (d) in Fig. 3 was measured.

This final characteristic may be considered in greater detail. It passes the zero point of inflection according to the dashed tangent with the slope $\tan\alpha_0$ indicating a zero resistance R_0 or a zero conductance G_0. At higher fields the electrostatic attraction causes the currents to rise over this initial tangent thus indicating a decrease in resistance or an increase in conductivity. This preliminary analysis picture will be completed later.

(b) Oscillograms

Because of the aforementioned phenomena, the static step-by-step method is somewhat inaccurate. An oscillographic representation is much more advantageous because it discloses not only the dynamic behavior of the electric oil in question, but, at the same time, assures greater stability because the chains assume an average state of polarization.

The simplest device for oscillographing non-linear I–V characteristics is indicated in Fig. 4. The colloidal resistor R is fed via the fixed preresistor r by the alternating voltage V_d of 60 cycles with variable amplitude. The cathode ray of the oscilloscope O—which, for the sake of simplicity, is shown without its amplifiers—is deflected horizontally by the voltage V_- across R, if the switch S is in "measuring" position, and vertically by the voltage drop $V_r = I \cdot r$.

Figure 5 shows a typical example of a non-linear I–V oscillogram (a). The straight calibration line (b) was obtained with the switch S in "calibration" position so that both inputs of the oscilloscope are in parallel. If Y and X are the deflecting sensitivities in

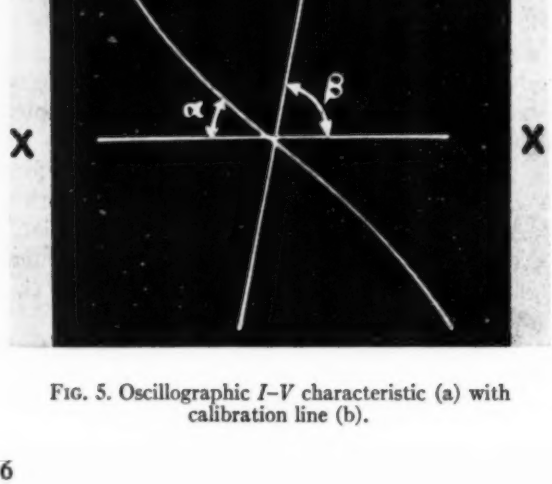


FIG. 5. Oscillographic I–V characteristic (a) with calibration line (b).

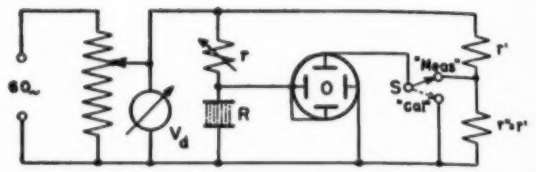


FIG. 6. Alternating current bridge for obtaining oscillographic bridge characteristics.

both axes, then their ratio is determined by the angle β: $Y/X = \tan\beta$. On the other hand, any angle α of an arbitrary chord connecting the point of any voltage V with the zero point or two arbitrary points at any voltages $\pm V$, whose tangent is proportional to the corresponding conductance $G = 1/R$, is given by the ratio of the horizontal deflection:

$$D_x = X \cdot V_d/(1+r/R)$$

and the vertical:

$$D_y = Y \cdot V_d/(1+R/r)$$

as:

$$\tan\alpha = D_y/D_x = (r/R)\ \tan\beta = r \cdot G \cdot \tan\beta. \qquad (2)$$

By means of this relation, R, G, or $\sigma = e/F \cdot R$ (F = surface of the electrodes) can be plotted *versus* V provided the horizontal deflection sensitivity (X) is determined by a simple voltage calibration.

A more sensitive device for oscillographing the non-linearity of a colloidal resistor is an alternating current bridge whose schematic is indicated in Fig. 6. It differs from Fig. 4 only in that the non-linear attenuator $r \cdots R$ is supplemented by the resistive attenuator $r' \cdots r''$, which balances the fundamental voltage drop across R so that only the non-linear deviations produce the vertical deflection.

A simple analysis as before yields:

$$\tan\alpha = \tan\beta \frac{(r'r/r''R)-1}{(r'/r'')+1}. \qquad (3a)$$

Naturally the bridge has its highest sensitivity when $r' = r''$ so that Eq. (3a) simplifies into:

$$\tan\alpha = \tfrac{1}{2}\ \tan\beta[(r/R)-1]. \qquad (3b)$$

The bridge operates at its maximum sensitivity if it is in balance at very low momentary voltages. This means that r must always be adjusted in such a manner that the middle portion of the oscillogram remains horizontal. Such an example is shown in Fig. 7 together with the calibration line (b) which is obtained as before with the switch "S" in "calibration" position.

In analogy to Fig. 3 the gradual formation of the chains can be demonstrated by increasing the alternating driving voltage V_d. Such an oscillogram is shown in Fig. 8 whose individual characteristics differ by driving voltages progressing in increments of five volts. As long as the colloidal resistor is in its virgin state, R is approximately infinite and a little deflection line

appears with the slope:

$$\tan\gamma \sim \tfrac{1}{2}\tan\beta.$$

As soon as the threshold field is reached, the oscillogram becomes non-linear and, after having passed a horizontal middle portion at $R=r$, simultaneously becomes erect and expands.

This turning, especially the middle portion, reveals a decrease of the zero resistance R_0. On the other hand, the bridge can easily be kept in balance by always adjusting r in such a manner that the turning-up is neutralized or that the non-linear bridge characteristic keeps a horizontal zero passage. The measuring points shown in Fig. 9 were obtained in this manner and will be explained in greater detail in Section 5.

(c) Analytical Characteristics

A complete analysis of the complex behavior of the colloidal resistors is not possible, at least not at the present time, because too many parameters are unknown and not computable. However, an analytical representation of the experimental I–V characteristics yields a better insight into the philosophy of the electric oil and electrical performance of any colloidal suspension. Further investigations may disclose the complex relations by which the individual parameters, i.e., the dielectric constant and conductivity of the particles and the carrier fluid, the form and size of the particles and their surface properties, the mixture ratio, and last but not least, the previous history are related.

The following analysis is based on numerous measurements of a graphite-oil suspension. The solid curve in Fig. 10 shows an example which has been averaged from numerous measuring points at a certain prepolarization and limited to a permissible current density. The characteristic is deduced by plotting the current density i per cm^2 instead of the entire current I, and the field strengths E, instead of the impressed measuring voltage V. Hence, Eq. (1a) is presented by the reduced form:

$$i=I/F=\sigma E=\sigma V/e \qquad (1b)$$

where F is the electrode area in cm^2, σ the conductivity in mhos per cm, and E the electric field strength in volts per cm. This reduction has an advantage in that the characteristic is referred directly to the conductivity as well as to the electric fields which are responsible for the dielectric chain formation and the resulting non-linearity.

It has been mentioned in Section 3(a) that the I–V characteristic passes its inflection point under the slope of the lower tangent:

$$\tan\alpha_0=(i_0/E_0)\tan\beta=\sigma_0\tan\beta. \qquad (4a)$$

The fact that σ_0 depends to a very large degree on the previous history, i.e., on the momentary polarization, makes any pre-calculation very difficult if not impossible.

The deviation of the solid curve from the initial zero tangent can be represented by a cubical parabola, at least at low polarizing fields, in the vicinity of the zero point. Hence, the lower portion of the actual i–E characteristic may be described by:

$$\begin{aligned} i&=\sigma_0E+\sigma_0KE^3 \\ &=\sigma_0E[1+KE^2]. \end{aligned} \qquad (5a)$$

The so-called "field coefficient" K characterizes the non-linearity in first-order approximation.

As soon as the polarizing field becomes increasingly stronger the characteristic remains below the simple parabola. This deviation can be perceived by diminishing K in relation to E:

$$K'=K/(1+PE^2), \qquad (6)$$

where P is an additional quantity which characterizes the non-linearity of the extended characteristic. Hence, Eq. (5a) assumes the more accurate form:

$$i=\sigma_0E[1+KE^2/(1+PE^2)]. \qquad (5b)$$

The evaluation of the measured characteristic in Fig. 10 gives $\sigma_0=2\cdot10^{-7}$ mhos per cm, $K=1.2\cdot10^{-5}$, and $P=9\cdot10^{-7}$.

By means of this analytical representation the condition can be extended to field strengths far above the permissible measuring range. Equation (5b), with the

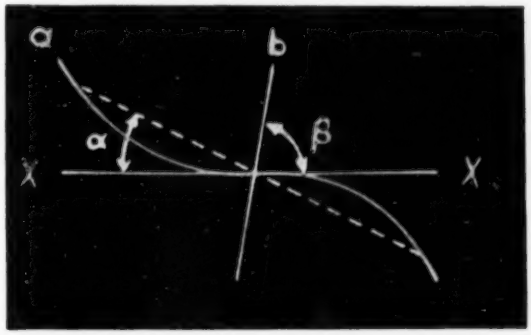


FIG. 7. (a) Non-linear bridge oscillogram with (b) calibration line.

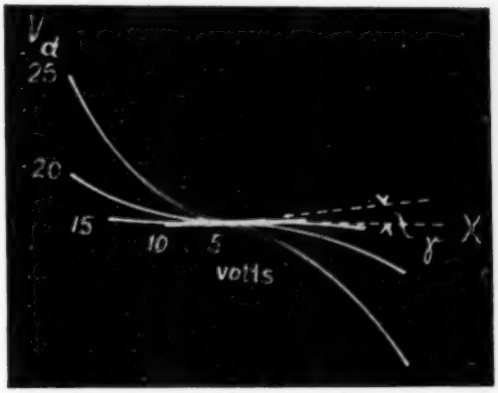


FIG. 8. Non-linear bridge oscillogram at successively increasing driving voltages.

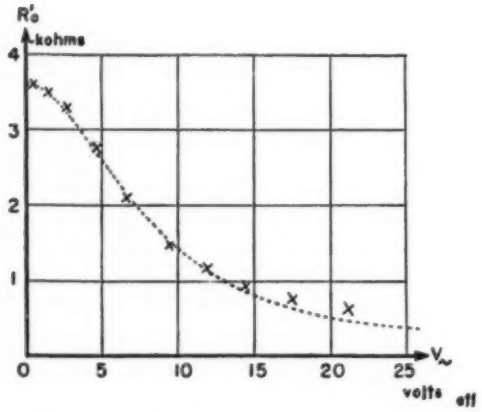


FIG. 9. The zero resistance R_0' in relation to the alternating voltage $V_\sim$ across the colloidal resistor.

experimentally determined values, yields the dashed characteristic shown in Fig. 10. Of special interest is the condition arising when it approaches the upper tangent whose slope follows when $E \to \infty$ as:

$$\sigma_{sat} = \sigma_0[1+K/P]. \tag{4b}$$

This saturation conductivity is the only quantity which can be understood very simply as characterizing the phenomenon that all semiconductive particles are compressed by infinitely powerful forces of electrostatic attraction in the form of a solid block. In other words, the saturation conductivity (in the present example 13 times the zero conductivity) must be identical with the conductivity of solid graphite, or, for example, coal, multiplied by the mixture ratio of the suspension, because the cube of unity is filled only in part by the graphite, while the rest is filled by the insulating carrier oil.

4. CONDUCTIVITY AND RESISTANCE CHARACTERISTICS

(a) Non-Linear Conductivity

The change in the relative direct current conductivity referred to zero conductivity can be computed easily from Eq. (5b):

$$\Delta\sigma/\sigma_0 = \sigma/\sigma_0 - 1 = K\cdot E^2/(1+P\cdot E^2). \tag{7a}$$

Using the values given above for K and P, this quantity is plotted *versus* E as the solid curve in Fig. 11. Since $P<K$ the curve, in the vicinity of the zero point, is a simple parabola:

$$\Delta\sigma/\sigma_0 = K\cdot E^2. \tag{7b}$$

At the field strength:

$$E_i = 1/(3P)^{\frac{1}{2}}, \tag{7c}$$

in the present example at 610 v per cm, a point of inflection appears around which the relative direct current conductivity changes linearly with E. At very high field strengths, the relative change in conductivity approaches, asymptotically, a saturation value:

$$(\Delta\sigma/\sigma_0)_{sat} = K/P, \tag{7d}$$

i.e., in the present example, 13.4.

If the conductivity characteristic should be computed from the non-linear bridge oscillogram, it can be done by means of the simple formula:

$$\Delta\sigma/\sigma_0 = 2\tan\alpha/\tan\beta.$$

Apart from the direct current conductivity, a differential or alternating current conductivity which is proportional to the slope of the I–E characteristic at any arbitrary operating point or at any arbitrary bias field E_B must be included. Differentiation of Eq. (5b) yields the relative differential conductivity, i.e., $\sigma_\sim$ referred to σ_0:

$$\sigma_\sim/\sigma_0 = 1+KE_B^2\frac{3+PE_B^2}{(1+PE_B^2)^2}. \tag{8a}$$

The dashed curve in Fig. 11 illustrates this relation between the differential conductivity and the polarizing field E_B. In contrast to the solid curve, the latter starts at 1, passes a flat maximum of constant conductivity,

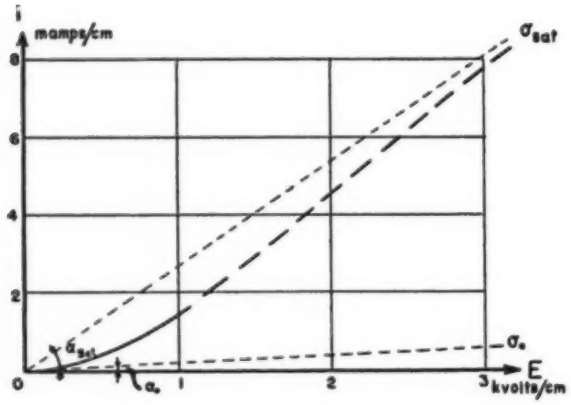


FIG. 10. Analytical I–E characteristic.

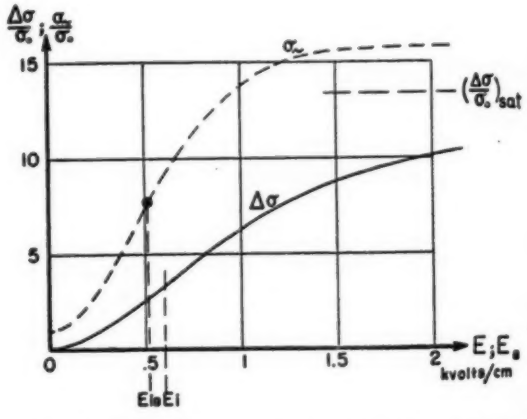


FIG. 11. The relative direct current conductivity $\Delta\sigma/\sigma_0$ and the alternating current conductivity $\sigma_\sim/\sigma_0$ *versus* the polarizing field strength E or E_B.

and approaches asymptotically the same saturation value as before.

At relatively low fields, in the vicinity of the zero point, Eq. (8a) simplifies to:

$$\sigma_{\sim}/\sigma_0 = 1 + 3KE_B^2 \tag{8b}$$

revealing a non-linearity three times greater than in $\Delta\sigma$.

This, at the same time, causes the inflection point to shift to:

$$E_{iB} \sim 1/2(P)^{\frac{1}{4}}, \tag{8c}$$

where the relative differential conductivity

$$(\sigma_{\sim}/\sigma_0)_i = 3KE_{\sim}/2(P)^{\frac{1}{4}}, \tag{8d}$$

again is proportional to any fluctuations of E superimposed upon the bias field E_{iB}.

(b) Resistance Characteristics

The non-linear conductivity yields the corresponding resistance characteristics. For reasons which will become more apparent later, it is convenient to introduce the differential resistance $R_{\sim}$ which follows from Eq. (8a) as:

$$R_{\sim} = \frac{R_0}{1 + KE^2[(3+PE^2)/(1+PE^2)^2]}. \tag{9a}$$

If it is considered in relation to the bias voltage V_B, the equation assumes the form:

$$R_{\sim} = \frac{R_0}{1 + KV^2[(3+PV^2)/(e+(P/e)V^2)^2]} \tag{9b}$$

or, if restricted to the permissible range of relatively low voltages:

$$R_{\sim} = \frac{R_0}{1 + 3KV^2/(e^2+2PV^2)}. \tag{9c}$$

Examination of this equation reveals that the voltage sensitivity of any colloidal resistor becomes higher the smaller the oil-gap, or, in first-order approximation, proportional to the second power of e.

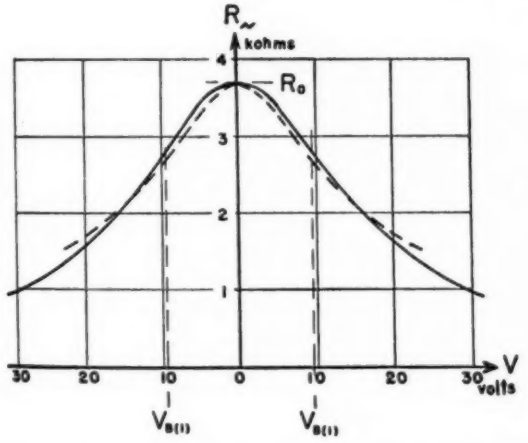


FIG. 12. Non-linear resistance characteristic of a colloidal test resistor with 1-mm electrode separation.

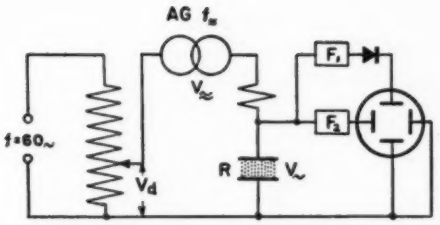


FIG. 13. Arrangement for obtaining oscillograms of differential resistance characteristics with the aid of an auxiliary-frequency f_z superimposed upon the oscillographing voltage $V_{\sim}$ of 60 cycles.

As an example, the solid curve in Fig. 12 illustrates the non-linear characteristic of a colloidal test resistor with $R_0 = 3.7$ kohms and $e = 0.1$ cm.

For the following experiments it is convenient to consider only two particular portions of the entire characteristic. First, the maximum portions around R_0 at zero bias which is simply described by:

$$R_{\sim(2)} = \frac{R_0}{1 + 3(K/e^2)V_{\sim}^2}. \tag{10a}$$

Second, the linear portions of both side branches around the bias voltage:

$$V_{B(1)} = e/3(K)^{\frac{1}{4}} \tag{10b}$$

in the present example 9.6 volts, at which operating point:

$$R_{\sim(1)} = -1.12R_0eV_{\sim}(K - P/3)/K^{\frac{1}{4}} \sim -1.12R_0V_{\sim}(K)^{\frac{1}{4}}/e. \tag{10c}$$

The negative sign does not indicate a negative alternating current resistance but signifies only that the resistance decreases if the momentary control voltage $V_{\sim}$ has the same polarity as the bias voltage.

Naturally, the advantage of a direct oscillography holds also for the resistance characteristics. In order to obtain such characteristics, the arrangement diagrammed in Fig. 13 has been developed to operate as follows.

The basic schematic resembles the previous Fig. 4 with the following modifications: The auxiliary-frequency generator AG superimposes an auxiliary-frequency voltage V_z of some kc upon V_d. The auxiliary-frequency drop across R, after being separated from the 60 cycles by means of the high pass filter F_1 and after rectification between the vertical amplifier output and its deflecting plate, deflects the cathode ray in the Y axis. On the other hand, the horizontal deflection is produced, similarly as before, by the 60-cycle voltage drop $V_{\sim}$ after the auxiliary-frequency is filtered out by the low pass filter F_2. In this manner the oscilloscope indicates the differential resistance $R_{\sim}$ in relation to the momentary voltage across R instead of the direct current resistance R. Note that $V_z < V_{\sim}$ so that only a very short portion of the I–V characteristic is controlled by the auxiliary-frequency.

An example of the differential resistance characteristic of a colloidal test resistor with one mm electrode separation is shown in Fig. 14. Replacing the colloidal resistor with various constant resistors introduces horizontal calibration lines, whereas the sensitivity of the X-deflection was calibrated by means of a known 60-cycle voltage. In this way the resistance oscillogram is completely calibrated. It is redrawn as the dashed curve into Fig. 12, indicating remarkable agreement

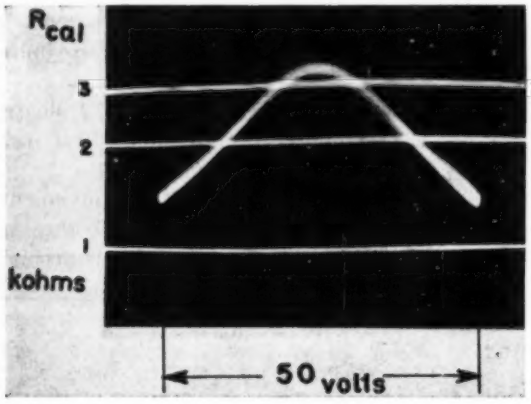


FIG. 14. Oscillographic characteristic of the differential resistance obtained at an auxiliary-frequency of 13 kc, with calibration lines.

with the analytical curve. As a preliminary result of these investigations it can be seen that the colloidal resistor has no drop in sensitivity with the auxiliary-frequency of 13 kc.

5. RECTIFICATION

(a) Pseudo Rectification

The aforementioned turning of the I–V characteristic under the influence of variable alternating voltages finds illustrative expression in the resistance characteristics. This is shown in Fig. 15 where three resistance characteristics are photographed successively at driving voltages $V_\sim$ increasing in increments of three volts. In analogy to Fig. 8 the resistance characteristics, at increasing $V_\sim$, shift downward.

According to these observations R_0 cannot be regarded as constant if the colloidal resistor operates under various voltage amplitudes but depends in turn on the state of polarization and, therefore, on the polarizing alternating field. This peculiar phenomenon can be included in the analysis in two different ways. Assuming that the dashed transfer characteristic in Fig. 16 is controlled by the alternating field $E_\sim = E_0 \cos\omega t$. Then, an average polarization takes place according to the effective field strength, i.e., proportional $E_{eff} = E_0/2^{\frac{1}{2}}$. As a result the initial tangent, whose slope characterizes σ_0, turns into the dotted chord which extends between the effective fields $\pm E_{eff}$ and whose slope indicates a new zero conductivity $\sigma_0' > \sigma_0$.

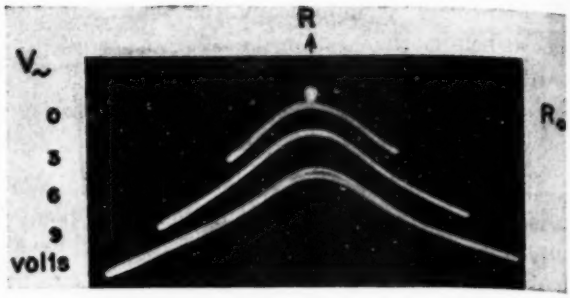


FIG. 15. Non-linear resistance oscillograms at various driving voltages $V_\sim$.

If for the sake of simplicity the simple Eq. (5a) is used, an analytical treatment of this average polarizing effect yields:

$$\sigma_0' = \sigma_0(1+(K/2)E_0^2) = \sigma_0(1+KE_{eff}^2), \qquad (11)$$

where σ_0 denotes the freely adjusting threshold conductivity just above the threshold field.

A more comprehensive explanation of the average polarization is given by considering the non-linear conductivity—or resistance—characteristic being controlled by a large voltage $V_\sim = V_0 \cos\omega t$. This is illustrated in Fig. 17. Here the fluctuating polarization causes a "pseudo-rectification," as shown on the right side of the figure. Because the chain formation is too inert to follow the alternating field itself, R_0 is diminished by the average value ΔR and, therefore, assumes the lower value $R_0' = R_0 - \Delta R$ where R_0 denotes, as before, the threshold resistance just above the threshold voltage. The analytical treatment of this resistive rectification yields, naturally, the same Eq. (11).

An exact example of the relation between R_0' and V_{eff} has already been shown in Fig. 9. There the dotted curve is computed by means of Eq. (11) in the form:

$$R_0' = 3.63/(1+1.51 \cdot V_{eff}^2 \cdot 10^{-2})$$

and is in agreement with the measuring points as well as could be expected.

(b) True Rectification

In addition to the resistive rectification which deals only with an average state of prepolarization, the curved I–V characteristic produces a true rectification. This is explained in greater detail by means of Fig. 18 where the non-linear diode characteristic is controlled by the alternating field $E_0 \cos\omega t$ at various operating points or at various bias fields E_B, respectively. Introducing $E = E_B + E_0 \cos\omega t$ into Eq. (5b) one obtains a rectified current density:

$$i_{rect} = \sigma_0 E_B \frac{1+(K+P)(E_B^2+E_0^2/2)}{1+P(E_B^2+E_0^2/2)}. \qquad (12a)$$

The dashed curves in Fig. 18 represent the true rectifi-

cation at two field amplitudes E_0. Obviously, the diode characteristic becomes erect as in Fig. 8 when subjected to an alternating control field. At very high fields $E_0 \gg E_B$, Eq. (12a) simplifies into:

$$i_{rect} = \sigma_0 E_B (1 + K/P). \qquad (12b)$$

As a result, the rectified diode characteristics straighten and finally approach the dotted line.

In order to prove this result by experiment, the non-linear bridge shown in Fig. 6 is supplemented, as illustrated in Fig. 19. The new arrangement differs from the simple bridge, in that an additional high frequency voltage $V_{\sim}$ of, e.g., 5 mc, is impressed upon the colloidal resistor. This high frequency is supplied by the signal generator "*SG*," whereby *R–F* chokes "*ch*" and capacitors C' prevent the high frequency energy from penetrating the bridge. Quite obviously the 60-cycle field E_{-} takes the place of the steady bias field E_B which produces, in connection with $E_{\sim}$, the "rectified diode characteristics" shown in Fig. 20.

Whether the true rectification is combined with the previously mentioned pseudo rectification or not, is open to conjecture. Be that as it may, the fact that the diode characteristic indicates no change in frequency up to the 50 mc presently attained, reveals a remarkably good frequency response. The frequency stability marks the colloidal resistor as a useful detector whose versatility will be revealed in greater detail by means of a subsequent paper.

6. NON-LINEAR LISSAJOUS-FIGURES, AND HARMONICS

Up to the present, all oscillograms have been considered only in reference to the characteristic of the colloidal resistor itself. The following section deals with the analysis of the electrical properties of a non-linear attenuator when the colloidal resistor operates in series with a constant resistor. For investigation from an electrical point of view, the non-linear bridge already shown in Fig. 6 is obtained, when combined with a linear attenuator.

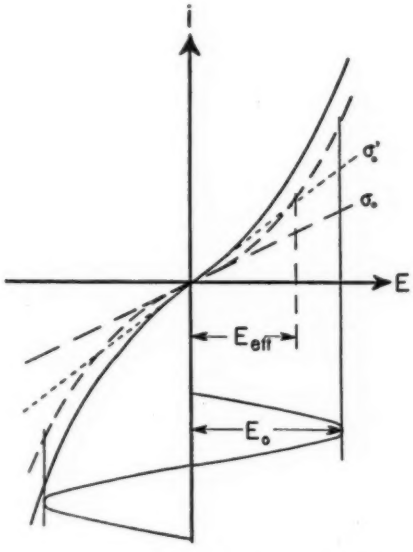


FIG. 16. Average polarization when under the influence of an alternating control field with the amplitude E_0.

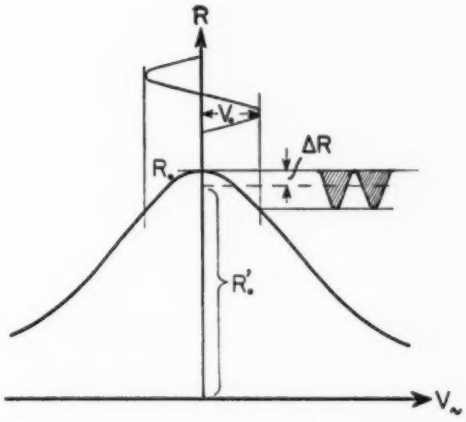


FIG. 17. Illustration of the pseudo rectification.

The foregoing bridge circuit, according to Fig. 21, is modified in that the horizontal input of the oscilloscope is reversed. The X-deflection is sinusoidal:

$$D_x = X V_d/2,$$

whereas the diagonal voltage across the middle branch of the bridge is:

$$V_{dg} = V_{-} - V_d/2 = (V_d/2)\{[2/(1+RG)] - 1\}.$$

Furthermore, an adjustable bias voltage V_B is added, insuring various operating points of the colloidal resistor.

For the sake of simplicity, the following analysis is restricted to two specific cases: First, the colloidal resistor operates under first-order non-linearity on linear branches of its conductance characteristic at a certain

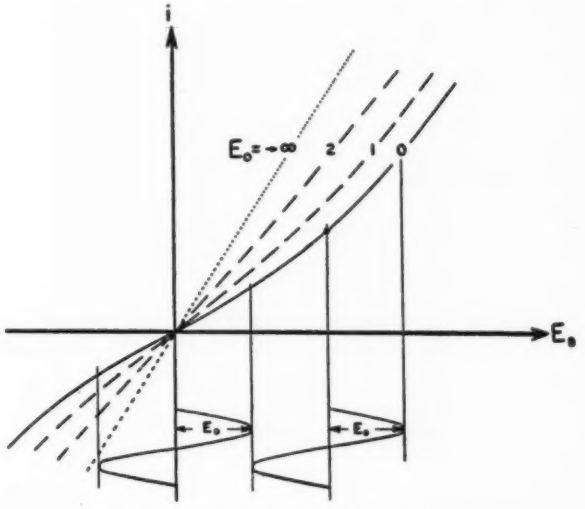


FIG. 18. "Rectified" diode characteristics.

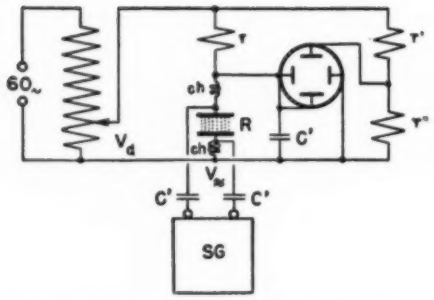


FIG. 19. Schematic of the non-linear bridge with superimposed high frequency voltages for oscillographing rectified diode characteristics.

bias voltage V_{Bi}; second, it operates under the non-linearity of second order, i.e., on the parabolic middle branch without bias.

In case one, the conductance characteristic is linear, following Eq. (7c) at the bias:

$$V_{Bi}=e/(3P)^{\frac{1}{2}}$$

with the slope:

$$G_B/V_d=\pm 3G_B eK/2(P)^{\frac{1}{2}}$$

so that the actual conductance at V_B becomes:

$$G_{\sim}=G_B[1\pm 3eKV_{\sim}/2(P)^{\frac{1}{2}}]$$

where G_B is the direct current conductance at the bias V_{Bi}. Introducing this expression into Eq. (13) and, if the bridge is balanced by means of $rG_B=1$, then:

$$V_{dg}=\frac{V_d}{2}\left[\frac{1}{1\pm eKV_{\sim}/3(P)^{\frac{1}{2}}}-1\right]. \qquad (14a)$$

A simple solution is possible only at very low driving voltages. Then the distortions of the voltage $V_{\sim}$ across R compared to the fundamental amplitude can be neglected. More accurately it means that only distortions of the first order should be considered as distortions caused by the fundamental voltage $V_{\sim}$ alone, and not distortions of the higher order caused by a cross-modulation. According to this simplification $V_{\sim}=V_d/2$. With this value, Eq. (14a) in first-order approximation becomes:

$$V_{dg}=eKV_d/12(P)^{\frac{1}{2}}$$

or furthermore, if V_d is introduced in the form $V_0 \sin\omega t$:

$$V_{dg}=eKV_0^2/24(P)^{\frac{1}{2}}\pm eKV_0^2/24(P)^{\frac{1}{2}}\cos 2\omega t$$
$$=\Delta V_{dg}\pm V_{dg\sim}. \qquad (14b)$$

Apart from a direct voltage ΔV_{dg}, corresponding to the aforementioned rectification, a second harmonic appears whose phase reverses with the polarity of the bias.

Plotting V_{dg} *versus* V_d produces a non-linear Lissajous-figure as shown in Fig. 22(a). If the polarity of the bias is reversed, the oscillogram inverts and assumes the shape (b) illustrating and thus proving the reversal of phase.

The Lissajous-figures are pure parabolas only if taken at low driving voltages. In the present examples, however, the driving voltage was relatively high, so that the parabolas are distorted, partly because the linear portions of the conductivity characteristic are overcontrolled, and partly because of the cross-modulation. The linear-time oscillogram shown in Fig. 22(c) reveals the additional distortion of the original second harmonic and indicates a full series of higher order harmonics predominantly of even order.

The second example of second-order non-linearity at zero bias yields, after introducing:

$$G_{\sim}=G_0(1+K(V_{\sim}/e)^2)$$

into Eq. (13), when limited again to low voltages, with $V_d=V_0 \sin\omega t$, and $V_{\sim}=V_d/2$:

$$V_{dg}=-3KV_0^3[\sin\omega t-(\sin 3\omega t)/3]/64e^2. \qquad (15)$$

Obviously the second-order non-linearity, characterized by the second power of the control voltage, causes a fundamental voltage to appear—because of true rectification—together with a harmonic of the second-plus-one, i.e., the third-order harmonic. Plotting V_{dg} *versus* V_d produces a Lissajous-figure similar to the non-linear bridge characteristic shown in Fig. 7. The fundamental can be suppressed by rebalancing the bridge by means of a slightly diminished r. At this rebalanced condition, only the third harmonic remains and the Lissajous-

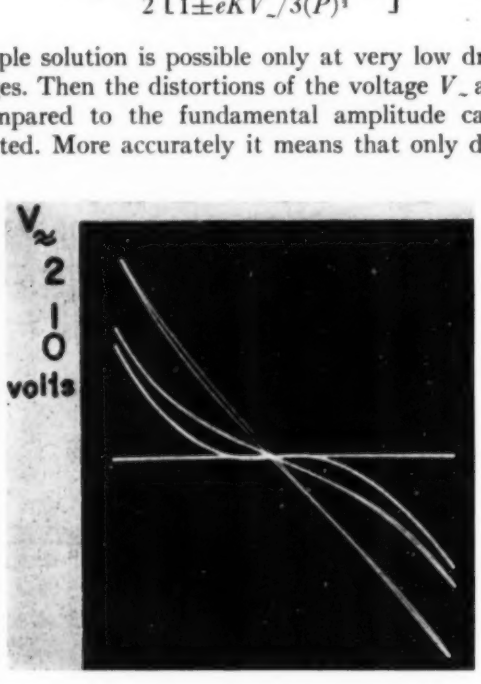


FIG. 20. Rectified diode characteristics obtained with an oscillographing frequency of 60 cycles superimposed by 5 mc.

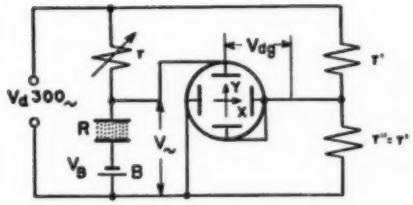
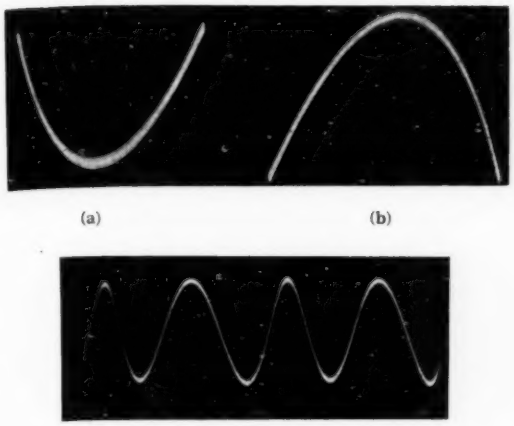


FIG. 21. Schematic of arrangement for producing the non-linear Lissajous figures shown in Figs. 22 and 23.



(a) (b)

(c)

FIG. 22. Non-linear Lissajous figures at first-order non-linearity.

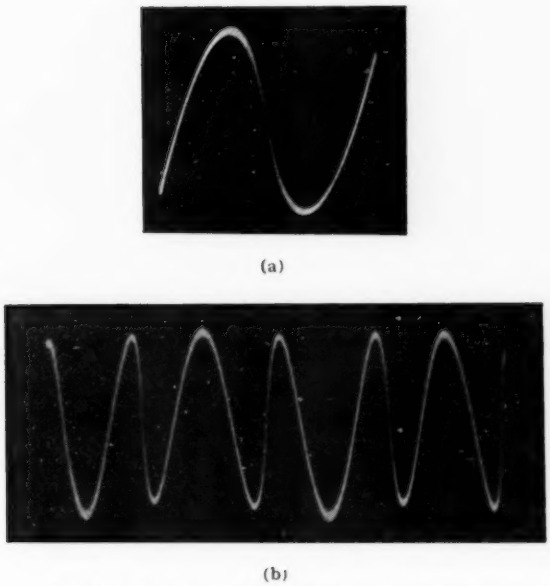
figure lies symmetrically on the X axis as shown in Fig. 23(a).

Here again, the oscillogram has been taken at relatively high driving voltages so that the validity of the developed formulas is exceeded. Consequently, the cross-modulation produces higher harmonics of an odd order as can be seen in the linear time oscillogram (b).

CONCLUSIONS

The concept of the formation of purely dielectric or of additionally semiconductive chains and the resulting non-linearities of dielectric constant or conductivity, respectively, may be important for understanding the philosophy of Seignette-electricity and some peculiarities of semiconductors.

On the other hand, the operational characteristics of the colloidal resistors assure certain interesting applications, as, for instance, new types of detectors, frequency converters and mixers, stabilizers and limiters for d.c. and a.c. voltages, linearizers for vacuum tube amplifiers, etc., where their large range of voltage sensitivity and power dissipation reveals entirely new possibilities. Furthermore, the sensitivity of the dielectric-semiconductive chains to mechanical or quasi-mechanical influences makes it possible to design sensitive conversion elements and transducers. In addition, when the dielectric-semiconductive particles are also ferromagnetic, magnetic forces of attraction can be superimposed upon the electrostatic attractions so that the resistance



(a)

(b)

FIG. 23. Non-linear Lissajous figures at zero bias.

of such a ferromagnetic resistor can be controlled in a manner similar to Hull's first magnetron. It remains for further experiments to prove whether or not it is possible to substitute the magnetic control field by an electric field and to develop liquid-filled amplifier and oscillator devices similar to the transistor.

During the preparation of this paper, it was possible to change from the liquid-filled colloidal resistors to dry types (the so-called polar resistors or "*Polaristors*") displaying basically all the properties mentioned before. All of these problems are under study and results of the investigations will be published in the near future.

ACKNOWLEDGMENTS

The author is indebted to the Acheson Colloids Corporation for furnishing him samples of Oildag, Castordag, and Glydag. Furthermore, the author wishes to extend his sincere thanks to Dr. Ernst O. H. Friederich for his efficient and able assistance in producing various samples of dielectric, semiconductive, and ferromagnetic materials. Last but not least the writer is very grateful to Mr. Harold Berman and Mr. Fred A. Koether for their valuable help in preparing this paper.

Internal Friction of Metals at Very High Temperatures*

T'ING-SUI KÊ
Institute for the Study of Metals, University of Chicago, Chicago, Illinois
(Received October 10, 1949)

In connection with the study of the internal friction peak (*versus* temperature) associated with the viscous behavior of grain boundaries in metals, some other effects were observed at higher temperatures causing an additional internal friction in superposition with the high temperature branch of this internal friction peak. This additional internal friction at high temperatures was found to have its origins in some effects of cold-working introduced in the interior of the grains which remains even after the complete recrystallization of the cold-worked specimens. It increases with the amount of cold-work the specimen was subjected to before and after recrystallization; it decreases with annealing at successively higher temperatures until a stable value is reached; and it increases with the precipitated impurity content in the specimen. This high temperature internal friction was found to be very high in an aged specimen of high purity aluminum alloyed with 4 percent of copper. These observations are consistent with the viewpoint that this internal friction is caused by the presence of dislocations in the interior of the specimen although the mechanism of giving rise to the internal friction is unknown. It is pointed out that a study of high temperature creep under very low stress may be conveniently carried out through the internal friction measurements described.

INTRODUCTION

IN the study of the high temperature internal friction of annealed polycrystalline aluminum, a peak has been observed when the internal friction is plotted as a function of temperature of measurement.[1] Such a peak was not observed in single crystal aluminum under similar conditions. Extensive study has shown that the internal friction peak is associated with the viscous behavior of the grain boundaries in metals. However, it has also been observed that the high temperature branch of this internal friction peak did not drop as rapidly as expected if this peak is entirely due to the viscous behavior of grain boundaries (Fig. 2 of reference 1). Apparently, some other effects occur at higher temperatures causing an additional internal friction in superposition with the high temperature branch of the grain boundary internal friction peak. The physical origin of this additional internal friction at very high temperatures is not understood.

The purpose of this paper is to report the results of a study on the factors influencing this additional internal friction. The physical origin of this internal friction will be discussed. All the internal friction measurements were made using a torsion pendulum as previously described.[1] The frequency of vibration was about one cycle per second. The applied stress was sufficiently small so that the magnitude of internal friction is independent of stress level at each temperature range concerned.

The additional internal friction at high temperatures has been observed in a number of metals in addition to aluminum. Figure 1 shows the internal friction peaks associated with the viscous behavior of grain boundaries for Mg,[1] Al,[1] Cu,[2] alpha-brass,[3] and Fe.[4] They were all measured in torsional vibration with a frequency of about one cycle per second. In each case it is seen that the internal friction increases again at temperatures higher than the optimum temperature of the grain boundary internal friction peak. Such a feature has also been observed by Köster in transverse vibration for Cu and Ag.[5] The frequency of vibration he used was 795 cycles per second for Cu and 642 cycles per second for Ag. It seems thus the occurrence of additional internal friction at high temperatures is independent of the mode of vibration and is a common feature of all metals.

This high temperature internal friction cannot be caused by the strain amplitude effect of the applied stress since, as we have mentioned above, the applied stress has been always kept small enough for each temperature so that the internal friction is independent of stress level. In making internal friction measurements with a torsion pendulum, a small load including the

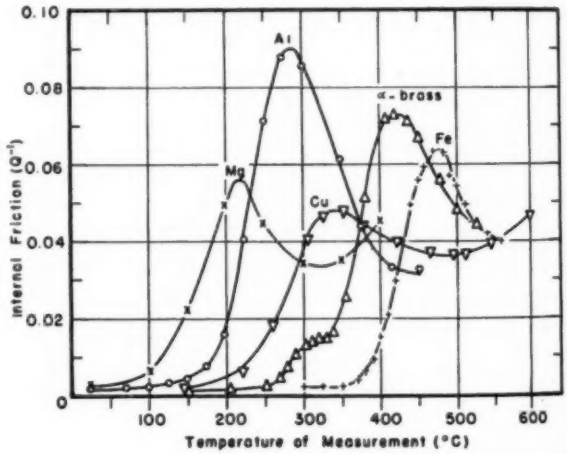


FIG. 1. Internal friction peaks (*versus* temperature) associated with the viscous behavior of grain boundaries in several metals.

* This research has been supported in part by ONR (contract No. N-6ori-20-IV, NR 019 302).

[1] T. S. Kê, Phys. Rev. **71**, 533 (1947).
[2] T. S. Kê, J. App. Phys. **20**, 274 (1949).
[3] T. S. Kê, J. App. Phys. **19**, 285 (1948).
[4] T. S. Kê, "Metals technology," Tech. Pub. 2370 (June, 1948).
[5] W. Köster, Z. Metallkde. **39**, 9 (1948).

inertia arm was applied longitudinally to the specimen. The effect of this longitudinal load has been tested for each metal. Figure 2 shows the internal friction for the case of iron when two different longitudinal loads were used in the measurement. It is seen that the whole curve is similar in both cases although one longitudinal load was seven times heavier than the other. Consequently, the additional internal friction at elevated temperatures is not attributable to the longitudinal load.

INFLUENCE OF COLD-WORK

We have mentioned previously that the high temperature internal friction is quite sensitive to the amount of cold-work the specimen was subjected to before recrystallization. Figure 3 shows the internal friction curves of specimens of recrystallized 99.991 aluminum: Curve I for the case when the specimen was subjected to a cold-work of 70 percent reduction in area and Curve II a cold-work of 95 percent reduction in area. Both specimens were recrystallized after being annealed at 450°C for five hours and the grain sizes of both were found to be similar. It is seen that the high temperature branch of the curve is lower in the case of the specimen subjected to a smaller amount of cold-work prior to recrystallization. This indicates that a smaller amount of prior cold-work causes a smaller addititional internal friction at high temperatures.† Curve III shows the internal friction curve of a specimen which was subjected to a small amount of cold-work after recrystallization. This cold-work has modified considerably the high temperature branch of the internal friction peak. It is apparent, thus, that cold-working both prior to and after recrystallization has the effect of raising the additional internal friction at elevated temperatures.

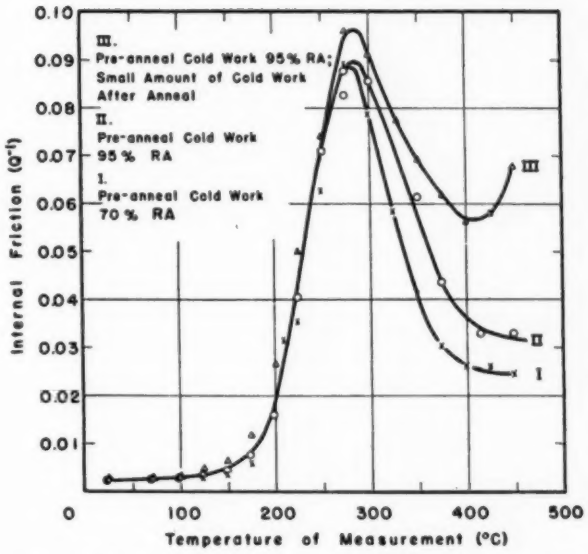


FIG. 3. Influence of the cold-work performed on the specimen before and after recrystallization.

This additional internal friction has been found to decrease as the annealing temperature goes successively higher. Figure 4 shows the effect of annealing on 99.991 aluminum subjected to a prior cold-work of 70 percent reduction in area. Figure 5 shows the case of aluminum of commercial purity (Al 2S) subjected to a prior cold-work of 96 percent reduction in area. It is to be pointed out that such a decrease of internal friction with an increase of annealing temperature reaches eventually a steady state which tendency can be seen by comparing the 400°C and 500°C annealing curves in Fig. 5.

ORIGIN OF THE HIGH TEMPERATURE INTERNAL FRICTION

The high temperature internal friction described above has also been observed in single crystal aluminum (99.6 percent purity) as shown by Curve IV in Fig. 6. The internal friction peak associated with the viscous behavior of the grain boundaries is absent in this case. However, the internal friction is seen to rise at very high temperatures. Curves I–III of Fig. 6 are the internal friction curves of several similar aluminum single crystal wires which were subjected, respectively, to the amount of cold-work shown and were then annealed at 450°C for one hour. The high temperature internal friction is seen to increase with the amount of cold-working. This indicates that this internal friction had originated from some effects of cold-working introduced inside the grains.

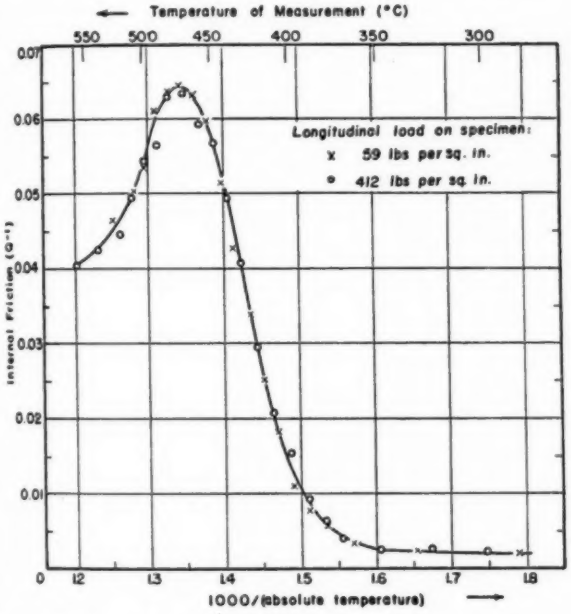

Temperature of Measurement (°C)
Longitudinal load on specimen:
x 59 lbs per sq. in.
o 412 lbs per sq. in.
Internal Friction (Q⁻¹)
1000/(absolute temperature)

FIG. 2. Effect of the longitudinal load applied to the specimen upon the high temperature internal friction.

† In the present case the position of the grain boundary internal friction peak did not shift with the different amount of prior cold-work (70 percent and 90 percent reduction). The situation is different when the amount of prior cold-work involved is smaller (see Fig. 2 of reference 2).

INFLUENCE OF IMPURITIES

A careful study of the internal friction curves shown in Fig. 1 revealed one interesting feature: The rise of internal friction at elevated temperatures is relatively higher when the grain boundary internal friction peak is relatively lower. Thus, in the case of aluminum, the grain boundary peak is the highest and the relative rise of the high temperature internal friction is the lowest. Previous experiments have shown that the viscous slip along grain boundaries can be partly or completely blocked by the presence of impurities at the grain boundaries. A lower grain boundary internal friction peak thus indicates a higher impurity content

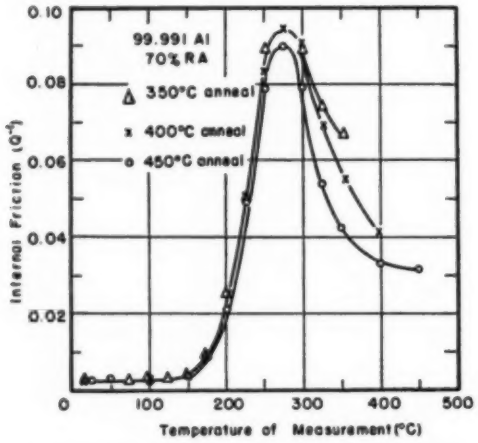


FIG. 4. Influence of annealing at successively higher temperatures (99.991 Al).

at the grain boundaries. The observation that the high temperature internal friction is larger whenever the grain boundary internal friction is smaller suggests that the former is connected with, or at least is influenced by, the presence of impurities.

Figure 7 shows the internal friction curves of polycrystalline aluminum of varying impurity content. Curve I is for high purity 99.991 percent aluminum; Curve II for aluminum 2S (containing as impurities 0.14 percent Cu, 0.48 percent Fe, 0.11 percent Si); Curve III for high purity aluminum alloyed with four percent of high purity copper. The prior cold-work and the subsequent annealing were similar in all cases. The grain sizes of the three specimens are also similar. The differences in these curves shown in Fig. 7 are thus attributable to the different amounts of impurity content. It is seen that in the case of aluminum containing 4 percent of copper, the grain boundary peak has been completely eliminated while the internal friction at 450°C is almost three times larger than the case when no copper was added.

STUDY OF AL-CU ALLOYS

In order to find out how the impurity content influences the internal friction at elevated temperatures, a detailed study has been made for the case of high

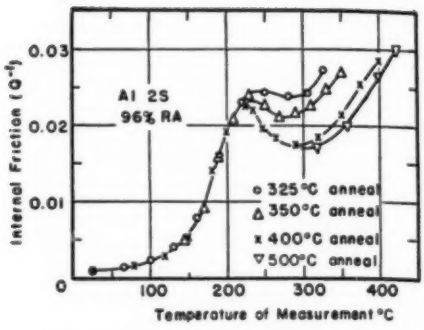


FIG. 5. Influence of annealing at successively higher temperatures (aluminum 2S).

purity aluminum alloyed with four percent of copper. In this case the impurity responsible for the effect concerned is undoubtedly copper. Figure 8 illustrates how the internal friction changes with heat treatment. The specimen was first solution-treated at 525°C for one hour and was then cooled slowly from 525°C so that the copper precipitates out in some form from the solid solution. It is seen that, after this treatment, the internal friction peak associated with the viscous behavior of the grain boundaries was eliminated while the additional internal friction at elevated temperatures is unusually high. The elimination of the grain boundary peak is supposedly due to the precipitation of copper in some form at the grain boundaries. As we have shown in the last section that the high temperature internal friction is not attributable to the grain boundaries, the rise of this internal friction is probably connected with the general precipitation of copper in aluminum, that is, the precipitation at the interior of the grains. The curve marked with (1) shows the case when the specimen was quenched from 525°C into water at room temperature. Under this treatment the

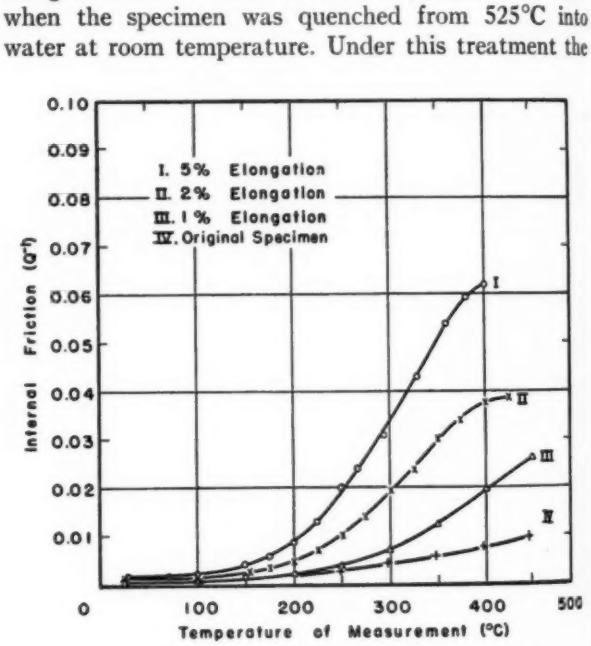


FIG. 6. High temperature internal friction of single crystal aluminum.

copper remains in solid solution in a supersaturated form. However, as the specimen is tempered at a higher temperature, the precipitation of copper will proceed at an increasing rate. The portion of curve marked (1) gives the grain boundary internal friction peak around 200°C which was measured by heating the specimen rapidly from room temperature to 200°C. This peak is fairly high although it is much lower than that for high purity aluminum without copper. This indicates that during the heating period some precipitation has taken place at the grain boundaries to block partly the viscous intercrystalline slip. When the specimen was heated rapidly from 200°C to 450°C, the internal friction at

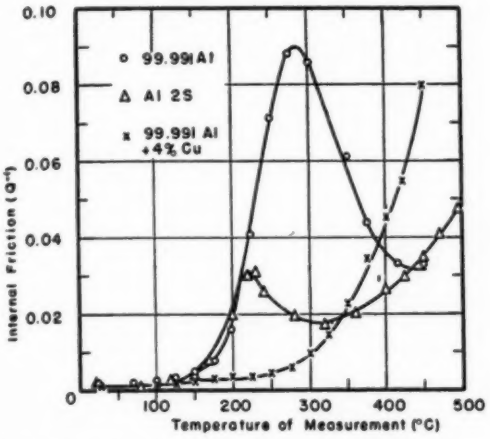


FIG. 7. Influence of impurity content upon the high temperature internal friction of aluminum.

450°C was measured and is given by (1). This internal friction was found to increase with time and its magnitude is shown by the time sequence of (1), (2), (3). This increase seems to be due to the fact that more copper was precipitated out from solid solution by tempering at 450°C. After the time sequence (3), the specimen was rapidly cooled from 450°C to 200°C. The internal friction at 200°C as given by (4) is much smaller than the original curve marked (1), indicating that more precipitation had been taking place at the grain boundaries during the tempering at 450°C. As the grain boundary precipitation is accompanied by general precipitation especially when the specimen is tempered at elevated temperatures, it is natural to assume that the rise of internal friction at 450°C by tempering is associated with the general precipitation of copper.

When the specimen was cooled slowly from 525°C, the solution treatment temperature, to 450°C, the precipitation of copper has probably been completed and the amount of copper left in solid solution is determined by the solubility of copper in aluminum at 450°C which is 2.60 percent by weight.[6] Under this condition the

[6] E. H. Dix, Jr. and H. H. Richardson, Trans. A.I.M.E. **73**, 560 (1926); Brown, Fink, and Hunter, Trans. A.I.M.E. **143**, 115 (1941).

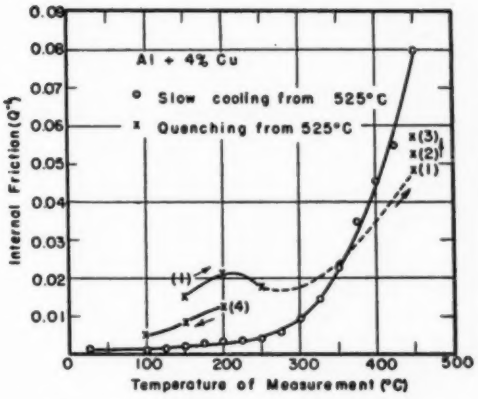


FIG. 8. Influence of heat treatment and the state of precipitation upon the high temperature internal friction of Al containing 4 percent of Cu.

internal friction at 450°C was found to be the highest as shown in Fig. 8. This indicates that the internal friction is the highest when the copper is in the precipitated state of some form.

The internal friction at elevated temperatures has been found to be independent of the impurities in dissolved form. Figure 9 shows the internal friction curves of high purity aluminum and of high purity aluminum alloyed with ½ percent of high purity copper. At 300°C the solubility of copper in aluminum is 0.45 percent.[6] The ½ percent copper in the specimen is thus in dissolved form at temperatures a little over 300°C. It is seen that the portions of the two curves over 300°C are almost identical, indicating that the ½ percent Cu in dissolved form has nothing to do with the internal friction at elevated temperatures. At temperatures lower than 300°C, the internal friction curve for the specimen containing ½ percent of copper is lower. This is evidently because of the blocking of grain boundary slip by copper precipitated in some form at the grain boundaries at temperatures lower than 300°C.

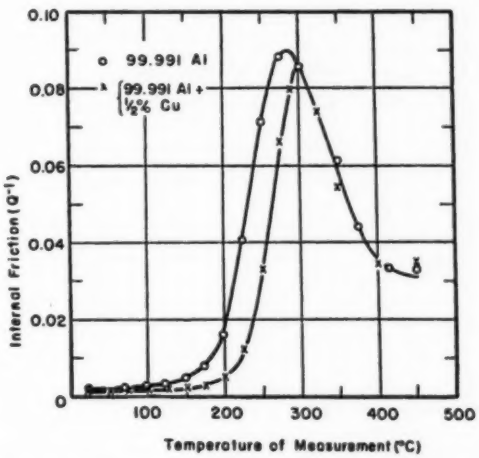


FIG. 9. Influence of dissolved impurity upon the high temperature internal friction.

TABLE I. Chemical and spectrographic analyses of aluminum specimens.

Alloying elements	99.991 Al	Aluminum 2S	Chemically pure Al
Cu	0.002%	0.14%	0.13%
Fe	0.002	0.48	0.053
Si	0.002	0.11	0.18
Mg	0.003		

STUDY OF TERNARY ALUMINUM ALLOYS

In describing the effect of impurity upon the internal friction at elevated temperatures, we have mentioned in connection with Fig. 7 that the high temperature internal friction of aluminum increases with the amount of copper contained in the specimen. Thus aluminum 2S containing 0.14 percent of copper has a higher internal friction than that of high purity aluminum but a lower internal friction than that of high purity aluminum alloyed with 4 percent of copper. A closer examination indicates that in the case of aluminum 2S, the situation is not so simple as the solubility of copper in aluminum at 450°C is 2.60 percent. The 0.14 percent of copper in the specimen should be in solid solution at 450°C and the high temperature internal friction should be small at this temperature. Figure 9 shows this is the case for high purity aluminum alloyed with ½ percent of copper. The anomalous condition in aluminum 2S indicates that a ternary or more complicated system is involved. The situation is more puzzling in the case of chemically pure aluminum containing 0.13 percent of copper, which gives an internal friction at 450°C as high as 0.11 (for a prior cold-work of 77 percent reduction in area. See Fig. 3 of reference 2). This internal friction is even higher than that of aluminum alloyed with 4 percent of copper at the same temperature. Table I gives the impurity content of the high purity aluminum and other samples.

It is seen that the copper content in aluminum 2S and in the chemically pure Al is about the same. The other impurities are chiefly iron and silicon. An insoluble Al-Cu-Fe phase or Al-Fe-Si phase may be formed and precipitates out from the solid solution although the copper and silicon, if present alone, will go into solid solution at the temperature range concerned. If this is the case, then the rise of the high temperature internal friction may be attributed to the presence of these ternary phases.

However, this still cannot explain why the internal friction of the chemically pure Al at 450°C is much higher than that of aluminum 2S. Figure 10 shows that at 450°C the internal friction of the former is higher than that of the latter although the previous cold-work in the former is only 42 percent reduction in area while it is 96 percent in the latter. It is not understood whether this difference can be attributed to the difference of the iron content in the two specimens (see Table I) or to some other unknown factors such as the past history of the specimens.

DISCUSSION OF RESULTS

In summary, we have found that the additional internal friction which dominates at elevated temperatures has the following features:

1. It occurs in polycrystalline and single crystal metals.
2. It increases with the amount of cold-work the specimen was subjected to before and after recrystallization.
3. It decreases with annealing at successively higher temperatures until a stable value is reached.
4. It increases with the precipitated impurity content in the specimen.

Let us examine now what are the underlying mechanisms of this high temperature internal friction.

A glance at Fig. 1 shows that the internal friction at elevated temperatures always occurs after the grain boundary relaxation has been completed. As the shear stress is relaxed along the grain boundaries, a concentration of stress will be built up at the grain corners.[7] Under proper conditions the local stress concentration may be large enough to initiate slip at the grain corners.[8] If this does happen it will certainly cause a considerable increase of internal friction. As a matter of fact, Hanson and Wheeler[9] have observed the macroscopic movement of grain boundaries at elevated temperatures. However, this was the case only when the applied stress is large. In our present case the applied stress is small, so the stress concentration at the grain corners is probably not large enough to initiate slip. This seems to be true as the curves in Fig. 7 show that the internal friction at 450°C is high in spite of the fact that the grain boundary relaxation was blocked by the presence of impurities.

If slip at grain corners has been initiated during measurements at 450°C, the internal friction should be sensitive to the stress amplitude. On the contrary, the internal friction has been found to be independent of stress levels under the experimental conditions.

In his study of amorphous plasticity, Orowan has pointed out that a disordered arrangement of atoms leads to a viscous type of flow if the temperature is

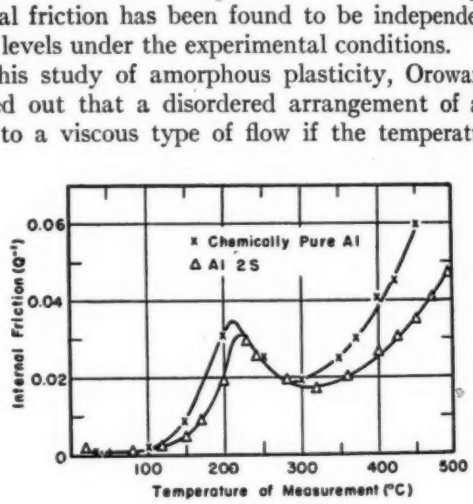


FIG. 10. High temperature internal friction of aluminum 2S and of chemically pure aluminum.

[7] C. Zener, *Elasticity and Anelasticity of Metals* (The University of Chicago Press, 1948), p. 127.
[8] C. Zener, Phys. Rev. **69**, 128 (1946).
[9] D. Hanson and M. A. Wheeler, J. Inst. Metals **45**, 229 (1931).

sufficiently high.[10] "Atomic disorder can be found not only at grain boundaries, but also inside the crystals if these have been plastically distorted, and particularly during recrystallization and phase transformations when the atoms leave one lattice and join another."[11] Along this line of thought this author has recently described a grain boundary model in which the transition region at the boundary was considered as consisting of numerous disordered groups of atoms or diffused holes.[2] The viscous intercrystalline slip was considered to occur through the atomic rearrangement by thermal agitation within each "disordered group" by a shear process involving as units of flow only a few atoms. Such a grain boundary model and slip mechanism are consistent with experimental facts. Among others, such a viscous intercrystalline slip will give rise to internal friction.

We have also pointed out that the "disordered groups" of atoms may conceivably exist also in the interior of the grains, except that they are much smaller in number than those existing at the grain boundaries. However, it can be shown that such disordered groups of atoms, when existing isolated in the interior of grains, cannot cause appreciable stress relaxation to account for the high internal friction observed at high temperatures as shown in Fig. 7 in which the internal friction of Al+4 percent of Cu is shown to be as high as 0.08 at 450°C. In the language of dislocation theory, we can say that numerous dislocations were introduced in the specimen during the cold-working and some dislocations remain therein even after the complete recrystallization of the specimen. The stress-induced movement of the dislocations can give rise to internal friction although the exact mechanism involved is not known.

The number of dislocations introduced in the specimen increases with the amount of cold-work. It is conceivable that the number of dislocations remaining in the specimen after recrystallization is larger in the case of a larger prior cold-work. The number of survival can be further reduced by annealing at a higher temperature until an equilibrium state is reached. In each case the internal friction increases with the number of dislocations existing in the specimen.

The effect of undissolved impurities on the observed internal friction may be primarily due to their retarding effect upon the removal of dislocations by annealing. This explains the observed facts that the high temperature internal friction is raised by undissolved impurities. However, additional consideration must be given to the complicated influences of impurities upon the stress-induced movement of dislocations.

CONCLUDING REMARKS AND CORRELATION WITH CREEP MEASUREMENTS

The foregoing discussions indicate that the experimental observations are consistent with the viewpoint that the internal friction at elevated temperatures under small stresses (of non-grain boundary origin) is caused by the presence of dislocations in the interior of the specimen. Accordingly, any factor that influences the number and the movement of dislocations in the specimen will result in a change of this internal friction. Such factors consist of the past thermal and mechanical treatments on the specimen, the impurity content, and the state of the impurities in the specimen.

Since a rise of internal friction has been shown to correspond to a rise of the rate of anelastic creep under low stress,[1,7] the additional internal friction at elevated temperatures may be correlated with the anelastic creep of metals at elevated temperatures in the same manner as the grain boundary internal friction was correlated with the anelastic creep of metals originating from the viscous behavior of grain boundaries (grain boundary creep).[1] Consequently, the discussions concerning the physical aspects and origin of the high temperature internal friction hold equally well for the case of anelastic creep under very small stress.

In conventional creep experiments the load applied to the specimen is usually large. This complicates the situation since the applied stress introduces additional creep properties to the specimen under test. To understand the physical aspects of the creep of the original specimen, the creep experiments must be done under sufficiently small stress so that the applied stress furnishes only a directive effect. A step toward this direction has been made by Chalmers in his study of the effect of applied stress upon the micro-creep of tin.[12] The study of creep through measurements of internal friction at low stress, as described in this paper, may prove to be both expedient and capable of giving unambiguous results for theoretical analysis.

[10] Ing. E. Orowan, Proc. West of Scotland Iron and Steel Inst. (February, 1947).
[11] R. Becker, Zeits. f. techn. Phys. **7**, 547 (1926).
[12] B. Chalmers, Proc. Roy. Soc. (A) **156**, 427 (1936); J. Inst. Metals **61**, 103 (1937).

The Orientation Dependence of the Rate of Grain Boundary Migration*

PAUL A. BECK, PHILIP R. SPERRY, AND HSUN HU
Department of Metallurgy, University of Notre Dame, Notre Dame, Indiana
(Received November 14, 1949)

Very small recrystallized grains of a large variety of orientations were produced locally in lightly rolled high purity aluminum crystals. The rate of growth of these small grains varied greatly, depending on their relative orientation with respect to the matrix crystal. Grains with an orientation corresponding to a [111] rotation of about 40° showed by far the highest rate of growth. These results indicate that recrystallization textures developing in the spontaneous recrystallization of deformed single crystals can be accounted for by the orientation dependence of the rate of growth.

THE large grains growing in an essentially strain-free fine-grained matrix of single orientation texture (coarsening) have quite definite orientation relationships with the matrix. In face-centered cubic metals, such as aluminum,[1] copper,[2,3] and nickel-iron,[4] the orientation relationship most frequently found corresponds to a rotation of about 30–45° around a [111] axis.

It was recently observed[1] that for the essentially strain-free grains growing in a highly oriented cold-worked matrix (rolled or compressed single crystal of aluminum), where the driving force of grain boundary migration is the strain energy stored in the cold-worked metal (recrystallization), the orientation relationship with the matrix is the same as for coarsening.[5] The identity of the orientation relationships in the two processes is all the more remarkable as the driving energy for recrystallization after large deformations is estimated to about 100 or 1000 times that for coarsening.

The understanding of these reorientation phenomena depends largely on the question whether the orientation relationship is determined by the availability of nuclei in definite orientations only, or whether it is determined by the dependence of the rate of growth on the orientation. Oriented nucleation was proposed by Burgers and Louwerse[6] for recrystallization and by Burgers[7] and recently by Kronberg and Wilson[3] for coarsening. Oriented growth in coarsening was suggested by van Arkel.[8] The possibility of interpreting the reorientation in recrystallization on the basis of an oriented growth theory, was pointed out by Barrett.[9] Recent work by three independent groups of investigators,[1,3,4] with three different metals, furnished overwhelming evidence to the effect that in coarsening "nucleated" by small grains, formed amidst the highly oriented matrix in a large variety of orientations by recrystallization after inhomogeneous local deformation (sawing or shearing of the edges, ball indentation, needle scratch), the grains growing to large size have the same orientation relationship to the matrix as the coarse grains formed spontaneously (without artificial nucleation). These experiments prove that the rate of growth in coarsening does depend on the orientation and that the orientations favored in spontaneous coarsening are actually ac-

FIG. 1. Strain induced boundary migration in high purity aluminum. Area swept by boundary while moving from 1 to 2, during last anneal, is essentially strain-free, has same orientation as cold-worked grain *A*. Oblique illumination. Magnification 75×.

* This work was supported by the ONR, U. S. Navy, Contract No. N6ori-165, T.O. 1.

[1] P. A. Beck and Hsun Hu, "Recrystallization texture and coarsening texture in high purity aluminum," Trans. A.I.M.E. **185**, 627 (1949).

[2] J. S. Bowles and W. Boas, "The effect of crystal arrangement on secondary recrystallization in metals," J. Inst. Metals **74**, 501 (1948).

[3] M. L. Kronberg and F. H. Wilson, "Secondary recrystallization in copper," Trans. A.I.M.E. **185**, 501 (1949).

[4] G. W. Rathenau and J. F. H. Custers, "Secondary recrystallization of face-centered Ni-Fe alloys," Philips Research Reports **4**, 241 (1949).

[5] Unpublished work, carried out at this laboratory, gave similar results for the recrystallization of rolled copper single crystals. Maddin, Mathewson, and Hibbard ["The origin of annealing twins in brass," Trans. A.I.M.E. **185**, 655 (1949)] found similar orientation relationship for the recrystallization of extended brass single crystals, and Kronberg and Wilson (see reference 3) for the recrystallization of lightly rolled polycrystalline copper with cube texture.

[6] W. G. Burgers and P. C. Louwerse, "Über den zusammenhang zwischen deformationsvorgang und rekristallisationstexture bei aluminium," Zeits. f. Physik **67**, 605 (1931).

[7] W. G. Burgers, "Rekristallisation, Verformter Zustand und Erholung" (1941).

[8] A. E. van Arkel, "Quelques phenomenes de recristallisation," Rev. d. Met. **33**, 197 (1936).

[9] C. S. Barrett, "Recrystallization texture in aluminum after compression," Trans. A.I.M.E. **137**, 128 (1940).

counted for by their high rate of growth.† It is significant that no observation was so far reported by any investigator that could not be readily based on oriented growth. Finally, Dunn reported[10] the variation of the rate of growth with orientation for coarsening in silicon ferrite. There has been, however, no similarly direct evidence available to prove that in recrystallization, too, the rate of growth does in fact depend on the orientation of the growing grain with regard to the matrix.

The chief difficulty in carrying out a recrystallization experiment, analogous to that described above for coarsening, has been the supposed impossibility of preventing spontaneous recrystallization. If the spontaneous formation of recrystallization nuclei could be prevented, even though nuclei artificially provided in various orientations were capable of growing, then the possible dependence of the rate of growth of these nuclei on their orientation could be studied.

It was recently found[11] that in high purity aluminum the spontaneous formation of recrystallization nuclei is rather rare up to deformations as high as 40 to 50 percent by rolling. Instead, essentially strain-free material is formed by migration of certain grain boundaries already present in the cold-worked material (Fig. 1). Starting with a relatively large initial grain size in order to minimize the effect of these boundary migrations, in the present experiments large grain-areas were made available for artificial nucleation. Artificial nucleation was effected by inscribing with a sharp needle one or two scratches into the surface of each large grain of the electrolytically polished portion of high purity aluminum specimens, which had been previously rolled 12 percent. After scratching, the specimens were annealed in salt bath for 1 hour at 350°C, and electrolytically etched to produce a thin oriented oxide layer, revealing variations in grain orientation.[12] Typical results are shown in Fig. 2. It is clear that in the immediate vicinity of the scratches many small grains were formed by recrystallization caused by the high local deformation. Of these small grains some served as nuclei for larger grains growing at the expense of the lightly (12 percent) deformed original grains. The relative orientation of the largest recrystallized grains with respect to the deformed matrix crystal was determined by the etch pit and optical goniometer method to an estimated accuracy of about ±1°. The results for 19 recrystallized grains found in 6 deformed matrix grains are shown in Fig. 3, with all matrix poles rotated to the same position, the "active" [111] axis being in each case in the center of the pole figure. The direction of rotation was reversed for some crystals so as to correspond in Fig. 3 to that of the other crystals. In all cases the relative orientation of the large recrystallized grains with respect to the matrix corresponded approximately to a [111] rotation of 40°. In some matrix grains two or even three [111] axes were found to be active. In Fig. 3 these are all separately rotated to the center of the pole figure.[13]

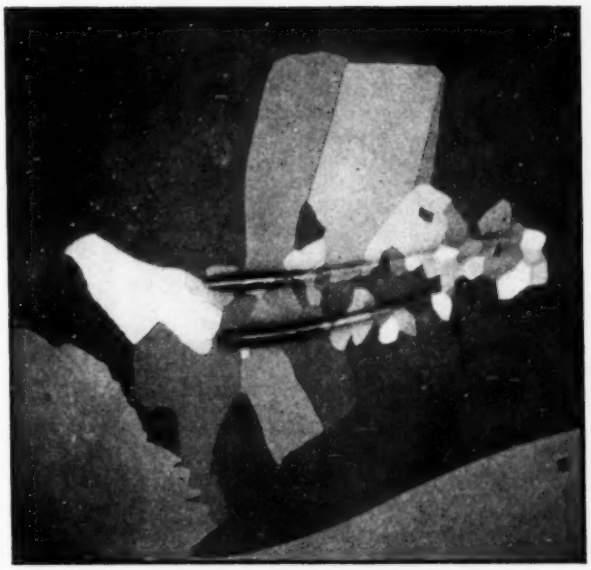


FIG. 2. Many small recrystallized grains of various orientations at the immediate vicinity of the two scratches. Selective growth of those recrystallized grains with suitable orientation with respect to the cold-worked matrix grain. High purity aluminum initially annealed 4 hr. at 650°C, rolled 12 percent, electrolytically polished and scratched, annealed 1 hr. at 350°C. Repolished and anodically etched. The two scratches were partially removed by repolishing in order to make the numerous small grains more clearly visible. Polarized light, Biot-Klein quartz plate. Magnification 25×.

† This conclusion was reached in references 1 and 4. Kronberg and Wilson (see reference 3) attempted to explain their results in terms of oriented nucleation.

[10] C. G. Dunn, "Recrystallization textures. Symposium on cold working of metals," Trans. A.S.M., 41A (1948).

[11] P. A. Beck and P. R. Sperry, "Strain induced grain boundary migration in high purity aluminum," J. App. Phys. **21**, 150 (1950).

[12] Paul Lacombe and Louis Beaujard, "Étude metallographique et cristallographique de la croissance et de la structure des pellicules d'oxydation anodique de l'aluminum," in the book published by the Comité Général D'Organisation des Industries Mecaniques, "*Etudes sur les Aspects des Pellicules D'Oxydation Anodique Formées sur L'Aluminum et ses Alliages*" (1944).
Andre Hone and E. C. Pearson, "Grain orientation in aluminum revealed by anodic film," Metal Progress **53**, 363 (1948).
P. R. Sperry, "Method for studying grain boundary migration in aluminum," Trans. A.I.M.E. **188**, 103 (1950).

[13] The fairly well-defined orientation relationship found here with artificial nucleation (Fig. 3) is in sharp contrast to the large scattering of orientations found by Burgers [W. G. Burgers and J. C. M. Basart, "Rekristallisation von Aluminium-Einkristallen I.," Zeits. f. Physik **51**, 545 (1928)] for the recrystallized grains in slightly deformed aluminum single crystals, without artificial nucleation. In accordance with a suggestion by F. C. Frank, the University of Bristol, England, it is very likely that the scatter of orientations in Burgers' experiments was a result of lack of nuclei in orientations most favorable for growth. Such a condition could have resulted from the small extent of the deformation used. In the present experiments nuclei of a sufficient variety of orientations were produced artificially. In the recrystallization experiment with copper, described by Kronberg and Wilson (see reference 3), the lightly deformed matrix was polycrystalline, with a cube texture. The sharp recrystallization texture found by these investigators was probably a result of the availability of a large variety of nucleus orientations, due to inhomogeneous deformation at grain boundaries.

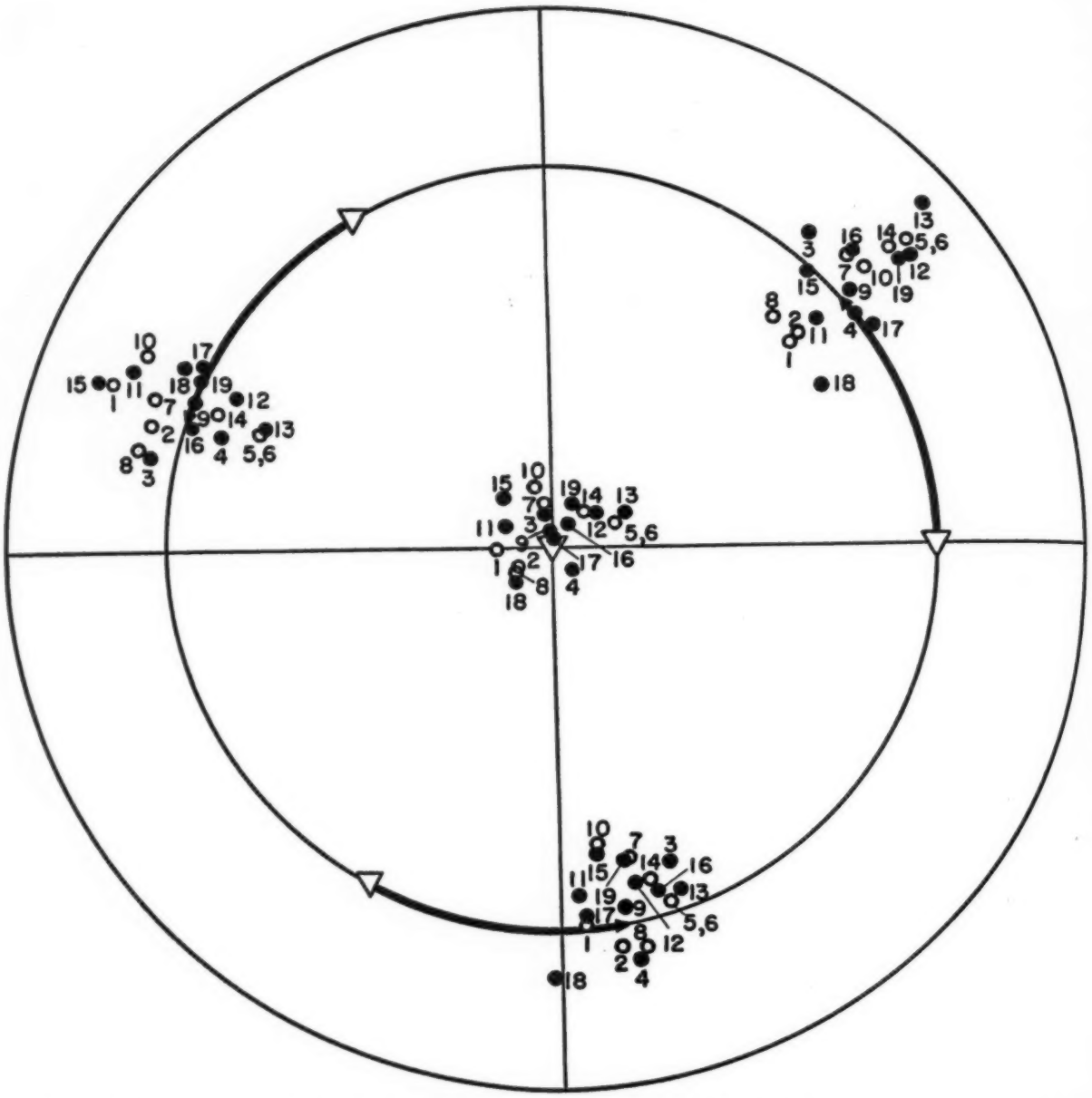


FIG. 3. (111) pole figure showing relative orientation of matrix grains △ (with the active [111] axis in the center) and of 19 largest recrystallized grains (circles). Arrow shows [111] rotation of 40°. Direction of rotation was reversed (by using mirror image of poles) for poles marked ●.

From the results shown it is evident that for recrystallization, as found previously for coarsening, the rate of growth of the artificially nucleated growing grains is highly dependent on their orientation relative to the matrix. In the recrystallization of high purity aluminum the crystal orientation which develops the highest rate of growth corresponds to a [111] rotation of about 40°–45° with respect to the matrix, as previously found for coarsening.[1] Since the same orientations were found[1] to predominate in recrystallization with spontaneous nucleation, it follows that the textures obtained in ordinary recrystallization (as well as in coarsening) can be accounted for by oriented growth, without recourse to oriented nucleation. It is interesting to note the high degree of independence of the orientation relationship between the fast-growing recrystallized grains and their matrices (Fig. 3) from variations in the orientation of the deformed matrix-grains with respect to the main directions of the rolling operation (Fig. 4). Another interesting observation may be made by inspection of Fig. 2. In certain directions the new grains grow much faster than in others; the rate of growth of the recrystallized grains is quite anisotropic. It appears that (111) faces tend to develop, suggesting that the rate of migration of such faces is lower than that of others.

Considering the controlling importance of the orien-

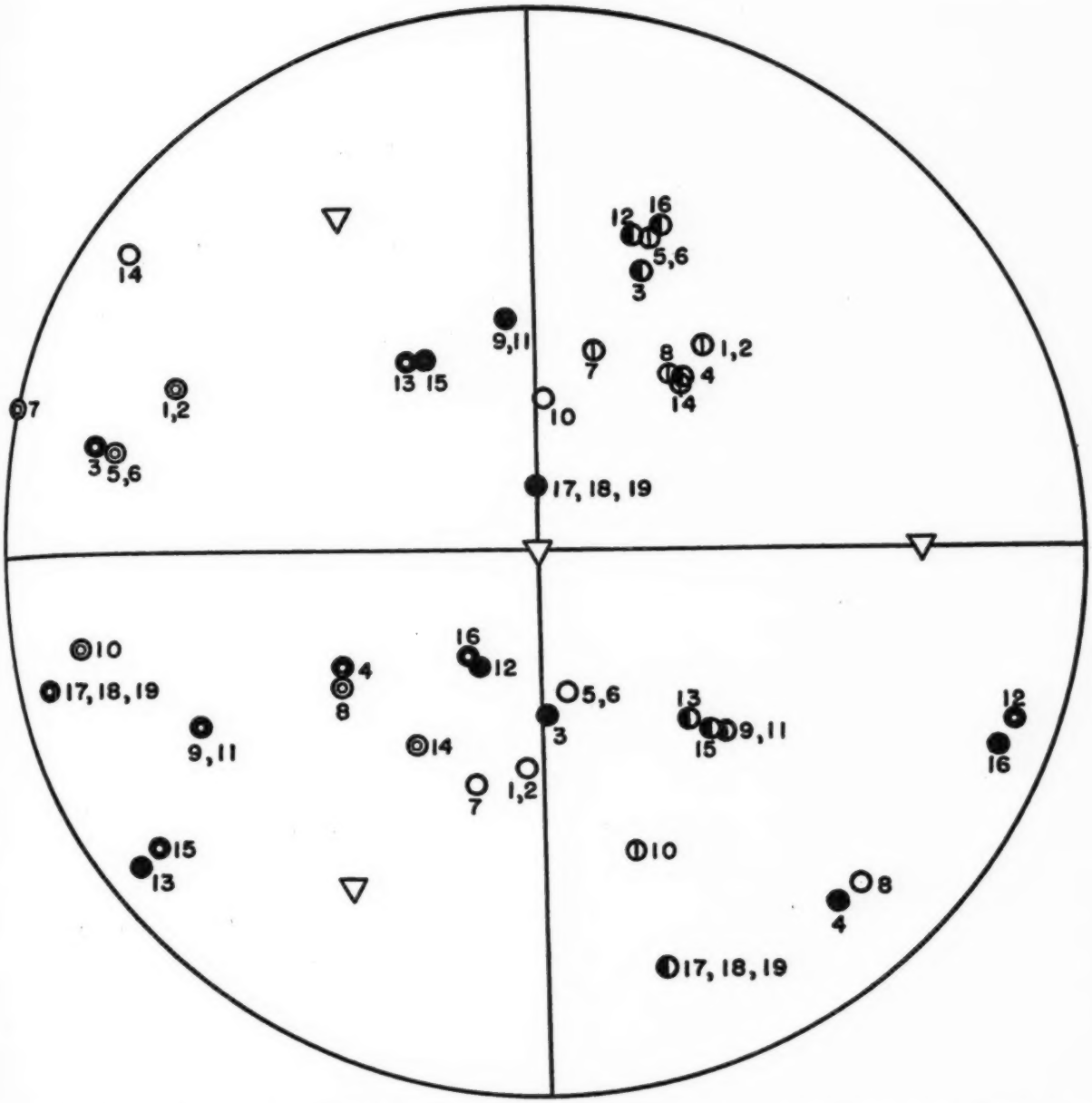


FIG. 4. Pole figure showing orientation of matrix grains △, with respect to the direction of rolling ○, the normal direction ● and the cross direction ◐ (active [111] axis in the center). The direction of rotation was reversed where circles are partially or completely filled.

tation-dependence of the rate of grain boundary migration, one would like to form a picture of this effect in terms of fundamental physical processes. In this connection, it is worth noting that the strain-free grains formed in cold-worked and annealed high purity aluminum in the manner shown in Fig. 1, are always growing at the expense of a strained grain of different orientation.[11] In observing a large number of such grains, they were *never found to grow at the expense of the strained grains of nearly the same orientation* with which they are in contact.[14] That the absence of migration of the boundary between a strained and a strain-free grain of nearly the same orientation is not due to

[14] It is interesting that, on the other hand, the sub-boundaries observed by P. Lacombe and L. Beaujard ["Les imperfections de structure des cristaux uniques d'aluminium pur," Revue de Metallurgie **45**, 317 (1948)] in polygonized aluminum are apparently highly mobile [A. Guinier and P. Lacombe, "L'état polygonisé, du cristal metallique," Métaux et Corrosion (1948), No. 212], even though they are located between lattice regions of nearly the same orientation. Possibly, this high mobility is connected with the fact that such a sub-boundary may be considered as a set of parallel line dislocations; the mechanism of boundary migration here is probably the one proposed by W. G. Burgers ["Recovery and recrystallization viewed as processes of dissolution and movement of dislocations," Proc. Koninklÿke Nederlandsche Akademie van Wetenschappen, **50**, 452 (1947)]. On the other hand, in the more general case, where two sets of intersecting edge dislocations are necessary to describe the boundary, it is more

FIG. 5. Strain induced boundary migration between grains *C* and *B* and between grains *A* and *B*. Boundary positions before and after the last anneal are marked 1 and 2. Strain-free area swept by moving boundary between *C* and *B* has orientation of *B*; it is growing into *C*. Strain-free area swept by migrating boundary between *A* and *B* has same orientation as *A*, and it is growing into *B*. High purity aluminum annealed at 650°C for 4 hr., rolled 12 percent, electrolytically polished, annealed 4 hr. at 350°C and 1 min. at 400°C, repolished mechanically and electrolytically and etched anodically. Oblique illumination. Magnification 75×.

a lack of driving energy[15] is obvious from the fact that the same strained grain is often seen being absorbed by a strain-free grain of a different orientation (Fig. 5). The surface energy connected with the migrating boundary (between two grains of different orientation) is undoubtedly larger than that connected with the stationary boundary (between two grains of practically the same orientation). Hence the energy required to increase the surface area of the boundary actually migrating is larger than the energy that would be required to expand the boundary which is in fact stationary; it so happens that strain induced boundary migration occurs just where the surface energy conditions are unfavorable for it. Clearly, then, the deciding factor must be something other than surface energy. It has been pointed out by Kronberg and Wilson[3] that the experimentally found relative orientations are near ones at which a certain percentage (such as 7 percent) of the atomic positions in the two lattices (that of the growing crystal and of the matrix) coincide, whereas generally in other orientations such coincidences are very rare. That an orientation relationship involving a large number of coincidences is not necessarily favorable for fast boundary movement is clear from the fact described above, that with nearly 100 percent coincidence the boundary is practically immobile. The fact that twin boundaries are particularly stable also points in the same direction. An exceptionally convincing demonstration of the lack of mobility of twin boundaries in the face of considerable driving forces was recently given by C. S. Smith.[16]

The experimental facts may be summarized by stating that with relative lattice orientations corresponding to a [111] rotation of 0° or 60° (twin orientation) the mobility of the grain boundaries is very low, and that with relative lattice orientations between these two, corresponding to [111] rotations of about 30°–45°, the mobility is the highest. Let us assume that grain boundary migration is a result of atoms leaving unfavorable positions in one crystal lattice, diffusing along the grain boundary and finally attaching themselves at some particularly favorable location to the crystal on the other side of the boundary. The necessity of *self-diffusion along the boundary* is an essential feature of the assumed mechanism of grain boundary migration. The effect of the relative orientation of the two crystals on the "mobility" of the boundary between them must be a result of its influence on the rate of transfer of atoms across the boundary. In view of the above mechanism of boundary migration, the most probable assumption as to the manner in which such an influence may be exercised is that the relative lattice orientation has a strong effect on the rate of self-diffusion along the grain boundary. It can be readily visualized for instance that at the boundary between two crystals of nearly the same orientation, or at a twin boundary, there is relatively little disturbance; the atoms are more tightly held in their positions, hence the rate of their diffusion is lower than along a boundary with more disorientation. Such an effect has apparently not as yet been found in diffusion studies, but its existence could be relatively easily ascertained by comparing the rate of self-diffusion along grain boundaries in a specimen of a single-orientation texture with that in a specimen of random orientation, other factors, such as grain size, being held constant.

CONCLUSION

(1) Evidence is presented to show that in slightly deformed high purity aluminum crystals the rate of growth of recrystallized grains of various orientations varies over wide limits.

(2) Those recrystallized grains which grow at the

likely that grain boundary migration takes place by individual atomic migration, as described further below.

[15] The fact that in fine grained high purity aluminum gradual grain growth is impeded in the presence of a strong single orientation texture [P. A. Beck and P. R. Sperry, "Effect of recrystallization texture on grain growth," Trans. A.I.M.E. 185, 240 (1949)] is probably closely related to the described effect. However, in this case the situation is not nearly as clear-cut as in the phenomenon just discussed, since the driving energy (in this case grain boundary surface energy) as well as the mobility becomes very small. The same applies to the remarkable stability of a sharply defined "cube texture" in copper (see reference 3), and probably also to the small "insular" grains left unabsorbed by a surrounding recrystallized grain of nearly the same orientation [P. Lacombe and Aurel Berghazen, "Relations d'orientation entre monocristaux métallique de recristallization et petits cristaux inclus," Métaux et Corrosion (1949), No. 281].

[16] Discussion by C. S. Smith of the paper by Hibbard, Liu, and Reiter, "Annealing twins in copper and 70–30 alpha-brass," Trans. A.I.M.E. 185, 635 (1949).

highest rate have relative orientations with respect to the matrix corresponding approximately to a 40° rotation around a common [111] axis.

(3) Since the orientation with the highest rate of growth closely approximates that found previously for the recrystallization texture of 70 to 80 percent rolled or compressed aluminum single crystals, it is concluded that the recrystallization texture can be accounted for by the orientation dependence of the rate of grain boundary migration. (Similar conclusion was reached earlier (see reference 1) regarding the coarsening texture.)

(4) The hypothesis is advanced that the orientation dependence of the rate of grain boundary migration is brought about by the dependence of the rate of self-diffusion along grain boundaries upon the relative orientation of the two grains meeting at the boundary. The assumed diffusion effect could be tested independently from its connection with grain boundary migration.

Measurement of the Thermal Diffusivity of a Soil by the Use of a Heat Pump

E. B. Penrod
Department of Mechanical Engineering, University of Kentucky, Lexington, Kentucky
(Received October 13, 1949)

An earth heat pump was set in operation to abstract heat from soil assumed to be in thermal equilibrium. Temperature-time graphs of the antifreeze and of the soil at distances of six, twelve, and eighteen inches below the center of the antifreeze line were plotted from experimental data. From these graphs data are obtained from which the thermal diffusivity of the soil is calculated. The soil was analyzed and found to fall in the general Casagrande classification of lean clay (*CL*) with low plasticity.

INTRODUCTION

IN designing the ground coil or earth heat exchanger for a heat pump system that uses the earth as its source of heat, the thermal properties of the soil and its general classification should be known.* When the earth heat pump[1] is operating on the heating cycle, for example, heat is absorbed from the soil by the cold fluid circulating through the earth heat exchanger. The time rate of heat flow through the earth depends partly on its moisture content which in turn depends on the grain size of the soil.

An original method is used to determine the thermal diffusivity of soils which may be applicable to geological and geophysical studies. The soil was analyzed and classified in addition to measuring its thermal diffusivity.

THE METHOD

After the earth heat exchanger had been buried in the soil a transient temperature distribution was created in the ground. This was accomplished by circulating a cold antifreeze solution continuously through the earth heat exchanger (Fig. 1) and the evaporator coils of the heat pump system. A section of the earth heat exchanger at station 3 (Fig. 2) was considered as an infinite line heat source when the heat pump was operating on the heating cycle.

It was assumed that the initial temperature of the soil is a function $f(r)$ of the distance from the antifreeze line and that the temperature at any time τ is a function $t(r, \tau)$. The differential equation for transient heat flow[2] in cylindrical coordinates is

$$\frac{\partial t}{\partial \tau}=\alpha\left(\frac{\partial^2 t}{\partial r^2}+\frac{1}{r}\frac{\partial t}{\partial r}\right). \tag{1}$$

Using finite differences[3] instead of differentials, Eq. (1) transforms to

$$\alpha=\frac{\Delta t/\Delta \tau}{(\Delta^2 t/\Delta r^2)+(1/r)(\Delta t/\Delta r)}, \tag{2}$$

where α is the thermal diffusivity of the soil.

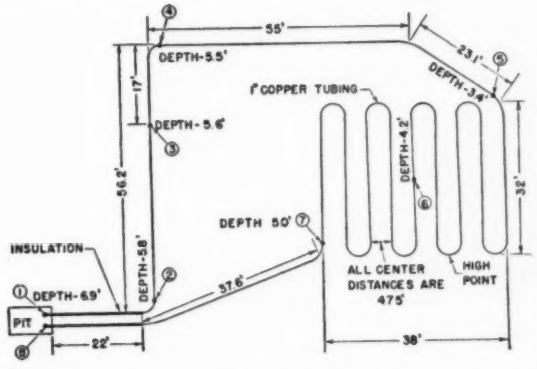


FIG. 1. Schematic diagram of the earth heat exchanger. The portion at station 3 was considered as an infinite line heat source.

* The Kentucky Utilities made the research possible by giving a Marvair heat pump and financial grant-in-aid to the University of Kentucky which supplemented the funds allocated by the Engineering Experiment Station of the University.

[1] E. B. Penrod, Am. Scientist **35**, 506 (1947).

[2] R. V. Churchill, *Fourier Series and Boundary Value Problems* (McGraw-Hill Book Company, Inc., New York, 1941), p. 165.

[3] Max Jakob, *Heat Transfer* (John Wiley and Sons, Inc., New York, 1949), Vol. I, p. 392.

TABLE I. Data for calculating $\Delta t/\Delta\tau$ and $\Delta t/\Delta r$.

Curve	Distance below center of antifreeze line, r ft.	Thermo-couple No.	Temperatures at noon, March 19, 1949, °F	Change in temperature from March 16 to 22, 1949, Δt°F	Time interval, $\Delta\tau$ hours
A	0.0	3	30.6	− 9.7	144
B	0.5	12	44.5	−10.9	144
C	1.0	13	48.6	− 9.0	144
D	1.5	14	50.1	− 5.0	144

The experiment was started on March 16, 1949 when temperatures of the antifreeze and soil were recorded at regular intervals for two days before the heat pump was put in operation. At 9:00 A.M., March 18, the plant was set in operation so that heat was absorbed from the soil as the antifreeze was circulated through the earth heat exchanger. While the heat pump was in operation temperatures were taken at regular intervals in the center of the antifreeze line (Fig. 2) at station 3 and in the soil below it at distances of six, twelve, and eighteen inches respectively. These data are shown graphically in Fig. 3.

Data for calculating the slope, $\Delta t/\Delta\tau$, of curves *A*, *B*, *C*, *D* at distances $r=0.0$, 0.5, 1.0, and 1.5 feet below the center of the antifreeze line are given in the last two columns in Table I. Curve *I* (Fig. 4) shows the variation of the slope with respect to distance below the center of the antifreeze line.

The temperature gradients, or first differences, were calculated from the data given in the fourth column of

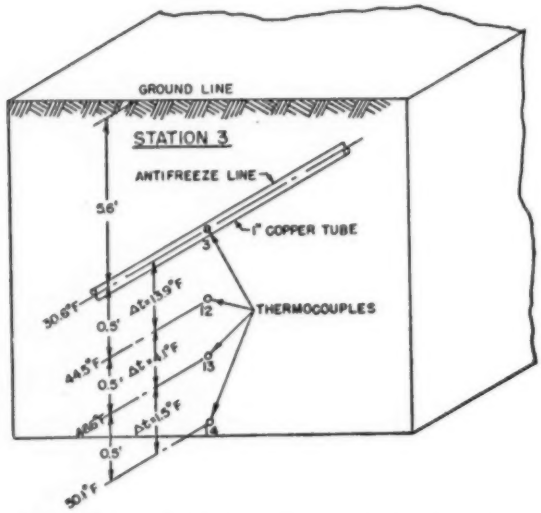


FIG. 2. Schematic diagram showing the location of the thermocouple beads at station 3.

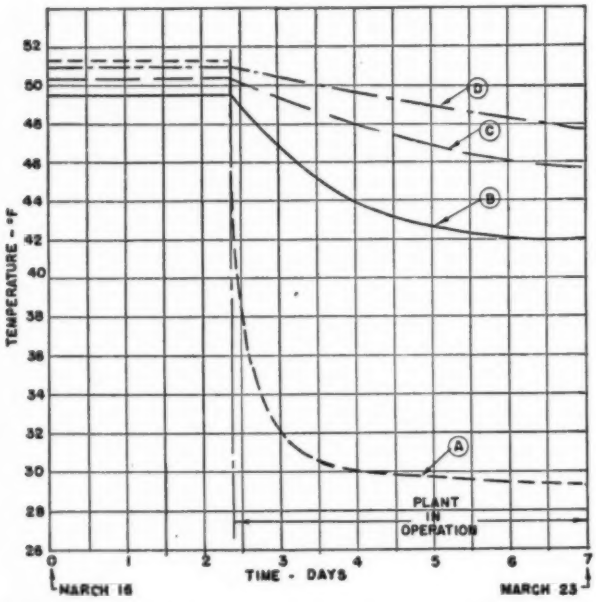


FIG. 3. Graphs showing the variation of temperature with respect to time at station 3.

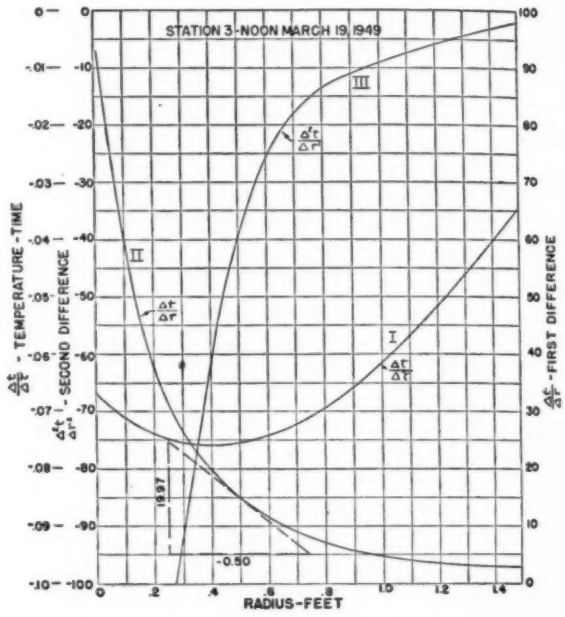


FIG. 4. Graphs used to determine the thermal diffusivity of the soil at station 3. Curve *I* was obtained from curves *A*, *B*, *C*, and *D* of Fig. 3 for noon, March 19. Curves *II* and *III* show the variation of the first and second differences with respect to the distance below the center of the antifreeze line.

Table I, and are listed in the third column of Table II. Curve *II* (Fig. 4) was obtained by coordinating values of the temperature gradient and distance from the center of the antifreeze line. The second differences, $(\Delta^2 t/\Delta r^2)$, were obtained by finding the slopes of curve *II* (Fig. 4) at various distances from the center of the antifreeze line. These results are given in the fourth column of Table II.

By substituting data given in Table II, in Eq. (2) values of the thermal diffusivity of the soil were obtained.

SOIL CLASSIFICATION

It is necessary to know the thermal properties of the soil as well as its general classification before an efficient

TABLE II. Calculated data and results.

Distance below center of antifreeze line, r ft.	Noon, March 19, 1949 $\Delta t/\Delta r$, °F/hour	$\Delta t/\Delta r$, °F/ft.	$\Delta^2 t/\Delta r^2$, °F/ft.²	Thermal diffusivity α, ft.²/hour
0.00	−0.0673			
0.25	−0.0752*	27.8	−115.0	0.0198
0.50	−0.0757	18.0	− 39.94	0.0192
0.75	−0.0715*	8.2	− 15.0	0.0176
1.00	−0.0625	5.6	− 9.0	0.0184
1.25	−0.0490*	3.0	− 5.0	0.0189
1.50	−0.0347			
			Average	0.019

* Values taken from curve *I*, Fig. 4.

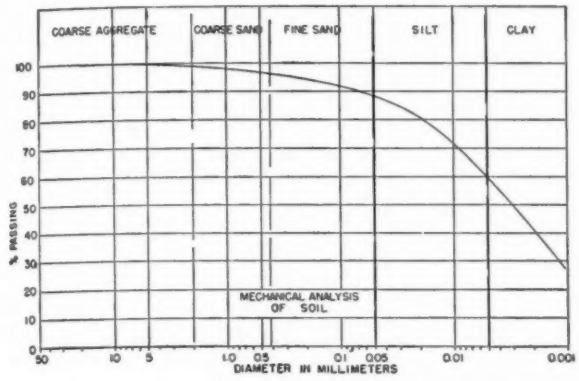


FIG. 5. Graph showing grain size of the soil in the vicinity of station 3.

earth heat exchanger can be designed. In order to determine the characteristics of the soil in which the earth heat exchanger was buried, a sample of the soil was taken at a depth of about 5 feet and 6 inches in the vicinity of station 3 at the time of the trench excavation.† Table III contains physical test data on the sample of soil, while results of the test for grain size distribution and plasticity characteristics are shown in Figs. 5 and 6. About 60 percent of its grains are in the size generally termed clay, and less than 30 percent of the sample is silt. The liquid limit of the sample is 43.4 percent and its corresponding plasticity index is 16.1 percent. The sample of soil falls in the general Casagrande classification of lean clay (*CL*) with low plasticity characteristics.[4]

By core sampling, the density of the soil in the vicinity of station 3 was found to be 120 pounds per cubic foot as received and 103.4 pounds per cubic foot after being dried to constant weight.

† The soil was analyzed by Messrs. L. E. Gregg and N. B. Drake of the Highway Materials Research Laboratory, Department of Highways of the Commonwealth of Kentucky.

TABLE III. Physical test data of the sample of soil analyzed.

Liquid Limit

Dish No.	No. of blows	Wt. of dish and wet soil	Wt. of dish and dry soil	Wt. of dish	Wt. of water	Wt. of dry soil	% of water	Liquid limit
66	29	9.01	8.00	5.56	1.01	2.44	41.4	43.4
29	24	9.98	8.63	5.56	1.35	3.07	44.0	

Plastic Limit and Plasticity Index

Dish No.	Wt. of dish and wet soil	Wt. of dish and dry soil	Wt. of dish	Wt. of water	Wt. of dry soil	% of water	Plastic limit	Plasticity index
72	6.26	6.10	5.55	0.16	0.55	29.1	27.3	16.1
142	5.94	5.84	5.45	0.10	0.39	25.6		

Casagrande classification—*CL*.

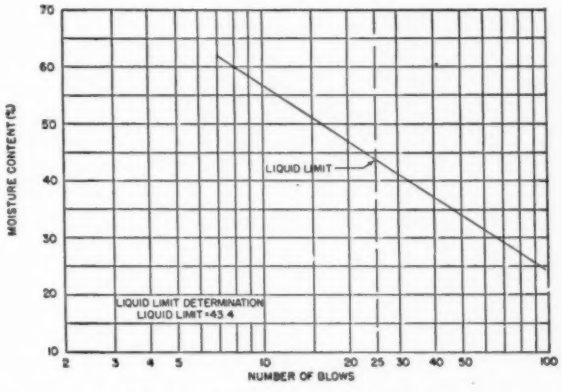


FIG. 6. Graph showing the moisture content of the soil in the vicinity of station 3.

CONCLUSIONS

The average value of the thermal diffusivity of the soil was found to be 0.019 ft.² hr.$^{-1}$ (0.0049 cm² sec.$^{-1}$) for the soil at a density of 120 pounds per cubic foot. The soil falls in the general Casagrande classification of lean clay (*CL*) with a low plasticity.

[4] "Classification and Identification of Soils," A.S.C.E. Trans. 113, 901 (1948).

I. Indoor Measurement of Microwave Antenna Radiation Patterns by Means of a Metal Lens

G. A. Woonton, R. B. Borts, and J. A. Carruthers
Department of Physics, McGill University, Montreal, P.Q.
(Received October 19, 1949)

The *E*-plane radiation patterns of two horn radiators have been measured in the laboratory with a metal lens by the same method that is used in optical diffraction pattern measurements. Comparison of patterns measured outdoors, by standard methods, with the lens measurements shows little deterioration in precision over the angular interval between ±15 degrees, for horns of aperture width 32 cm, measured by a lens of aperture 110 cm, at a wave-length of 3.2 cm. It was found that a stepped lens was unsatisfactory for this purpose, possibly because of diffraction caused by the steps.

INTRODUCTION

THE radiation pattern of an antenna is a statement of the power that the antenna abstracts from a plane electromagnetic field, expressed as a function of its angular position in the field. The pattern usually takes the form of a graph between power, relative to maximum received power, and angle of rotation measured from the point of maximum reception. Microwave antenna patterns are measured by rotating the antenna on a turntable, in a field set up by a transmitter some tens, hundreds or thousands of feet away. It has been shown by Cutler, King, and Kock[1] that the pattern produced in this way, approximates the true pattern to the extent that the spherical wave front from the transmitter approximates a plane wave in both phase and amplitude. It follows from Eq. (11) of their paper that, as the wave-length is decreased, it becomes increasingly more difficult to find a suitable, clear site for antenna pattern measurements, or conversely, that for a given site there is a wave-length limit below which the precision of measurement deteriorates.

Experiments have shown that, under favorable conditions, antenna patterns can be measured in the laboratory, with the help of a lens, by the same method as that used in optical diffraction measurements. Since both the precision and the convenience of making measurements of this type may be expected to increase with decreasing wave-length, it appears, that in these optical-type measurements there may be found a technique which will supplement standard radio methods in the wave-length range where a small outdoor site tends to become useless. The method is as follows: The spherical wave front from a point source is made plane for pattern measurement, by a metal lens, to an approximation that is determined by incidental diffraction from the aperture of the lens. Subject to restrictions on wave-length, on the sizes of the antenna and the lens, and on the angle of rotation, errors in the laboratory measurement due to incidental diffraction can be calculated and the calculations can be used to correct the measured pattern to the true plane wave pattern. The theory of these calculations, for linear radiators, has been discussed by G. A. Woonton.[2]

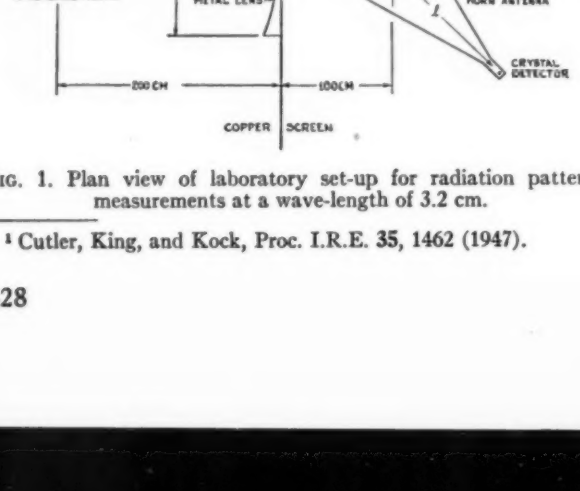


FIG. 1. Plan view of laboratory set-up for radiation pattern measurements at a wave-length of 3.2 cm.

EXPERIMENTAL METHODS AND RESULTS

Figure 1 is a schematic plan view of the measuring apparatus. Figure 2 is a photograph giving some of the details of the equipment. All measurements were made at a wave-length of 3.2 cm.

The source of radiation was the open end of a piece of 1 in.×½ in. wave guide which had been terminated with a plane flange. The open end was fed with 3.2 cm energy, through a wave meter and a calibrated antenuator section, from a Type 723A/B Klystron. The open end of the guide was placed about five feet above the laboratory floor, in line with the center point of the metal lens and at the focal point which was 200 cm behind the back surface of the lens.

Two metal lenses were used in these experiments, both of which were designed, from formulas given by Kock,[3] to have a focal length of 200 cm. The plate spacing for both lenses was chosen to be 2 cm which resulted in an effective refractive index of 0.6, at a wave-length of 3.2 cm. The first of these was a stepped lens of square aperture, 114 cm to the side, and was designed to focus in all planes. The plate spacing was maintained by five columns of metal spacers. It was found that antenna patterns, measured with this lens, exhibited unpredictable divergences from the true

[1] Cutler, King, and Kock, Proc. I.R.E. **35**, 1462 (1947).

[2] G. A. Woonton (to be published).

[3] W. E. Kock, Proc. I.R.E. **34**, 828 (1946).

radiation pattern and, after a series of modifications, the lens was discarded. The evidence is not conclusive, but it is believed that the final failure of this lens is associated with diffraction from its steps. During the tests with this lens it was found that the metal plate-spacers, which were $\frac{3}{16}$-in. metal cylinders, contributed unwanted diffraction, as did substitute Bakelite spacers. This trouble was eliminated by spacing the plates with acetate foam blocks. The performance of the lens was critically sensitive to the transverse position of the source. The condition was remedied by shielding most of the lens aperture so that the radiation passed through a rectangular slit 110 cm long by 22 cm high, with its long dimension in the plane of measurement.

The second lens was designed for the measurement of linear radiators only, and has been found satisfactory; the measurements which appear in Figs. 3, 4, and 5 were made with it. The lens contour is an elliptic cylinder which focuses in the *E*-plane. The aperture is a rectangle 114 cm in the *E*-plane by 20.6 cm in the *H*-plane. The plates are spaced by slabs of Styrofoam which completely fill the inter-plate space. The lens was mounted in a wooden frame with its center plate parallel to and five feet above the floor. Energy from the source was prevented from reaching the antenna, by paths other than through the lens, by means of a copper screen approximately nine feet square around the lens. The front, uncurved surface of the lens fitted snugly against the edge of an aperture (110 cm by 20.6 cm) which was cut in this screen; the edges of the aperture were given sharp definition by soldering the copper screen around a frame made of formed brass rod.

The correct position of the lens and source and the general adequacy of the set-up were determined by comparing the measured field in front of the aperture with the predicted diffraction field; simple optical theory, applied to the case of a plane wave, of length 3.2 cm incident on an aperture 110 cm by 20.6 cm was used in the calculation. Figure 3 shows that, by careful adjustment, the measured field can be made to approximate the predicted field. In the course of these measurements, it was learned that strong reflections from the windows and the walls could be minimized by setting the lens at an angle to the reflecting surface. The field measurements were made with a dipole and crystal detector mounted on an optical bench. The auxilliary apparatus was the same as that used for antenna measurements.

The *E*-plane patterns of the two horn radiators have been measured by this lens method; Figs. 4 and 5 give the results of these measurements. The field in which the horns were measured was square-wave modulated at six hundred cycles per second by modulating the reflector voltage of the klystron. A crystal detector attached to the apex of the horn converted the modulated received field into an audio voltage, which was amplified by a tuned audio amplifier and measured by a diode and direct current microammeter. The system was calibrated, before the measurements, by plotting the direct current output from the diode against the attenuation setting of a precision, calibrated, vane-type attenuator, which had been interposed between the klystron and the wave guide mouth. The antenna pattern was constructed by recording relative power in decibels (as determined by the microammeter and this graph) against the angular position of the horn in the field.

The copper horns on which experimental measurements were made were flared linearly in both the *E*- and *H*-planes from the end of a piece of standard, laboratory, *X*-band wave guide to a square aperture 32 cm to the side. Figure 4 presents pattern data on one of these horns, which was approximately 50 cm long; Fig. 5 presents data on the other which was about 100 cm long. Results from three different experiments are compared on each of these graphs: the solid line is the horn pattern, measured at an outdoor site, over a measuring path of about 225 feet, following standard procedure; it is considered to be a close approximation to the true antenna pattern. The dashed line is the mutilated distant pattern; a rectangular aperture, of the same dimensions and orientation as that used in the laboratory with the lens, was cut in the center of a large copper screen. The screen and its aperture were erected, at the outdoor site, one meter in front of the horn, between the horn and the distant measuring point 225 feet away. The horn pattern was then measured, following standard procedure. The pattern of the horn, as measured in the lens field in the laboratory, is indicated by crosses.

CONCLUSION

The radiation pattern of a linear antenna can be measured in the laboratory by optical methods, to a precision that is set only by incidental diffraction from the aperture of the lens. If the lens aperture is several times larger than the antenna aperture, the contribution

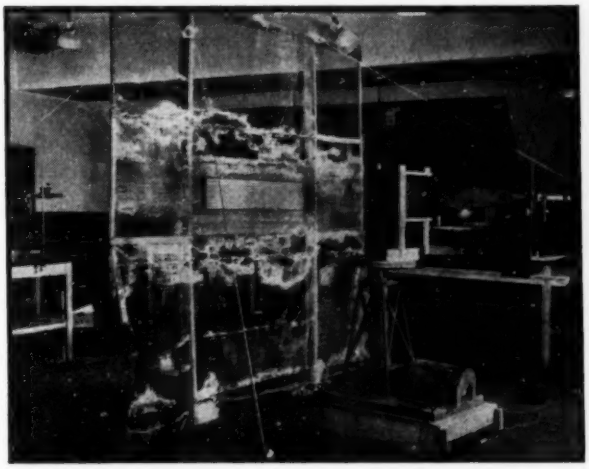


FIG. 2. Equipment used in laboratory measurement of antenna patterns.

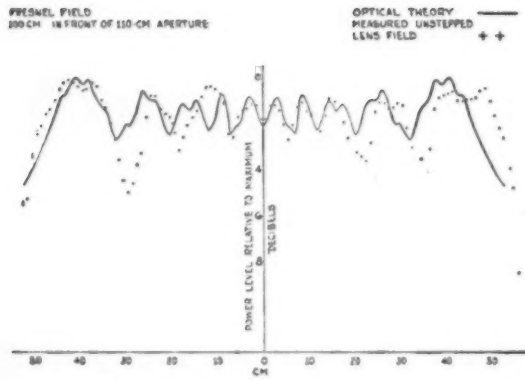


FIG. 3. Comparison of measured and predicted Fresnel field at a distance of 100 cm in front of the lens.

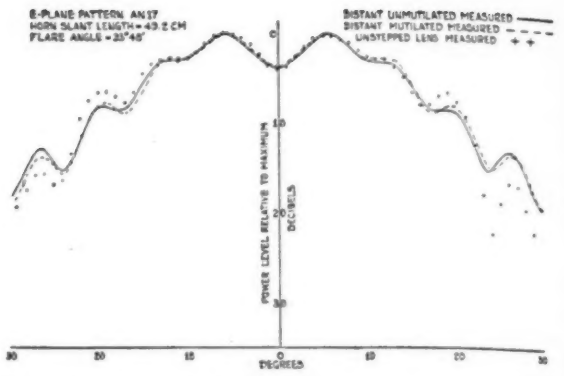


FIG. 4. Comparison of the true radiation pattern, the mutilated distant radiation pattern, and the radiation pattern measured through the lens for the *E*-plane of the 50-cm horn.

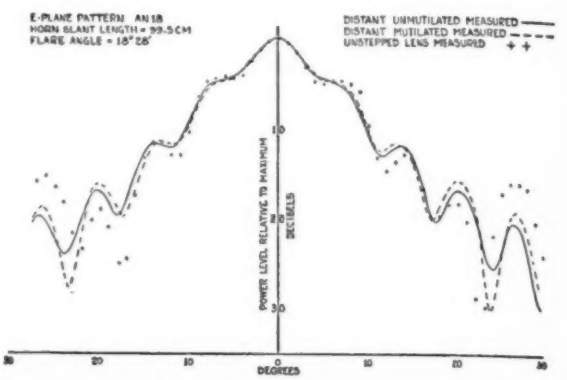


FIG. 5. Comparison of the true radiation pattern, the mutilated distant radiation pattern and the radiation pattern measured through the lens for the *E*-plane of the 100-cm horn.

of incidental diffraction to the center region of the pattern is negligible. Good agreement between patterns measured by the lens method and by the standard outdoor technique was obtained over the angular interval $-15° \leq \alpha \leq 15°$ for a lens aperture of 110 cm, an antenna aperture of 32 cm, and at a wave-length of 3.2 cm; beyond $\pm 15°$ the measured positions of the side lobes are approximately correct, but the relative amplitudes depart from those found by outdoor measurements. The effect of room reflections on these measurements can be minimized by proper location of the apparatus and when so minimized, reflections are not a serious source of error. A stepped lens has been found to be unsatisfactory in this application, probably because of diffraction caused by the steps and in general it has been found necessary to take more care in the design and construction of a lens for measuring purposes than has seemed essential with those to be used as antennas. Although it was found necessary to limit these measurements to the *E*-plane of the horn radiators, it is believed that there is no reason, other than physical size, why the method cannot be applied to all planes of an antenna of any kind.

ACKNOWLEDGMENT

This research was made possible by a grant from the Defence Research Board of the Department of National Defence, Ottawa, Canada. The stepped lens was built and some of the first experiments on it were performed by Messrs. J. H. Blackwell and H. M. Dutton with the senior writer, at the University of Western Ontario. Funds for the construction of the stepped lens and other apparatus used at that time were supplied from a grant made by the National Research Council, Ottawa, Canada. One of us (R.B.B.) is indebted to the National Research Council for a Bursary during the winter term of 1948–49.

Load Induced X-Ray Line Broadening in Nylon Filaments

JOHN KAUFFMAN* AND WALLER GEORGE*
Naval Research Laboratory, Washington, D. C.
(Received October 21, 1949)

The load induced microscopic structural changes in highly oriented Nylon filaments were followed by changes in the x-ray fiber diffraction pattern as a function of the resulting strain. Observations made included (1) changes in the dimensions of the fiber spots, and (2) a decrease in ratio of the intensity of amorphous to crystalline x-ray scattering. These new results for Type 200 du Pont Nylon indicate a change in the micro deformation mechanism at a strain corresponding to the inflection of the load-strain diagram (14 percent elongation). The mechanism for smaller plastic strains appears to involve a rotation of the crystallized domains into alignment with the fiber axis, accompanied by a small over-all additional crystallization. For strains larger than the inflectional strain, the mechanism seems to involve a progressive reduction in the domain size, and a reduction in alignment with the fiber axis. Both effects appear to saturate in the vicinity of the rupture strain. These results are discussed in terms of load induced effects found in metal single and polycrystals. It is suggested that the qualitative features observed may be representative of a wide class of load or strain induced micro deformation reactions in the solid state.

INTRODUCTION

THIS paper** describes a portion of a general research directed toward an understanding of the micro mechanisms of plastic flowing in crystalline high polymers. X-ray diffraction was the tool used for the study, and the material chosen was a strong, highly oriented, crystalline Nylon fiber obtained from E. I. du Pont de Nemours, Inc., denoted as Type 200. The fibers were received as 100 denier monofilaments.

The fine structure of oriented Nylon approaches the state found in polycrystalline metals, but in the case of crystallized polyamides, the crystallites are composed of long, complex molecules instead of atoms. These linear molecules exhibit a high degree of geometric organization within the region of the crystallite. The boundaries between crystallites, while imperfect and ill-defined, consist of what is essentially an "amorphous" material composed of molecules in a disordered state, which may be pictured as coiled and interlaced without geometric simplicity.

The microstructural state of 100 denier filaments of Type 200 is indicated by the x-ray fiber patterns as shown in Fig. 1. They are qualitatively similar to those obtained from polyamide fibers by others.[1] These were obtained using Ni filtered Cu$K\alpha$-radiation at room temperature with film specimen distance of 6 cm. The spots and arcs in the pattern result from periodicities in the crystalline material which scatter the radiation coherently. The disordered intercrystalline material produces the poorly defined amorphous halo at about the reflecting angle of the two very prominent equatorial spots. It is possible to obtain an estimate of the ratio of amorphous to crystalline material present in the specimen by comparing the intensity of the crystalline scattering to the amorphous scattering.[2]

EXPERIMENTAL PROCEDURE

The structure of the material as it is indicated by diffraction patterns was followed as a function of the amount of plastic strain from the initial condition of the specimen through tensile strains up to the rupture point, for strains resulting from a constant time rate of load application of 2.6 grams/denier/minute.

The specimens were in the form of a bundle of 25 monofilaments, and were deformed in tension in a commercial Scott IP-4 inclined plane testing machine.

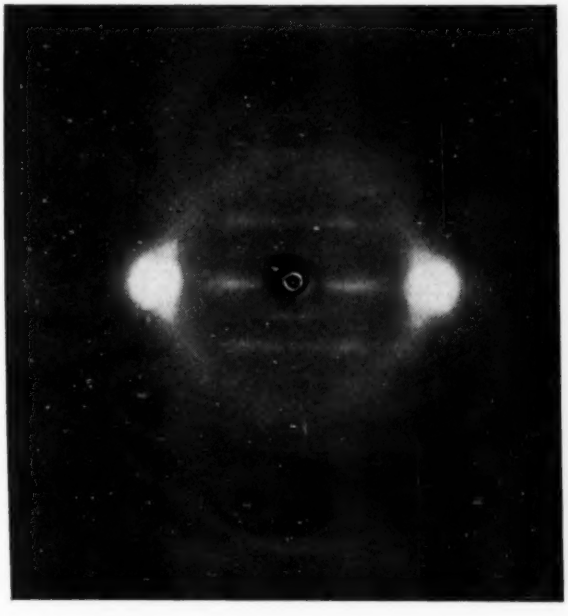


FIG. 1. Typical fiber pattern from Nylon bundle as received. Exposure *ca.* 1 hour. Film-specimen distance 4 cm Cu$K\alpha$-radiation.

* Members of the Staff, Mechanics Division.

** The new results reported in this paper comprise a portion of a thesis submitted by John W. Kauffman to the graduate faculty of the University of Maryland in partial fulfillment of the requirements for the M.S. degree (June, 1949). In abbreviated form, the results were discussed at the Cambridge meeting of the American Physical Society [Phys. Rev. **76**, 467 (1949)].

[1] I. Fankuchen and H. Mark, J. App. Phys. **15**, 364 (1944).

[2] W. Weidinger and P. H. Hermans, J. App. Phys. **19**, 491 (1948).

The machine used was equipped with "drum"-type specimen clamps which had been highly polished to reduce the friction at the fiber-drum interface. As the specimens reached various strains, the loading was abruptly terminated, the specimen clamped at a constant strain, removed from the testing machine, and the diffraction pattern immediately taken. In order to detect effects which might have an appreciable time dependence, additional patterns were taken after the specimen had been under a constant strain for several days; however, no such effects were observed for these times. Hence, one may consider long-time relaxation effects to be negligible for these experiments and therefore it is possible to relate the structural changes shown by the diffraction pattern to the points on the load-elongation curve at which the loading of the specimen was terminated. The time elapsed from the termination of the loading cycle until the beginning of radiation was approximately thirty minutes and held essentially constant for all experiments. This time interval apparently is sufficiently long to allow completion of short-time relaxation processes in all cases. This latter point is demonstrated in Figs. 2a and 2b, which show patterns obtained within one hour after the termination of the loading epoch and forty-eight hours, respectively. The strain used in this case was almost the strain at rupture, 26 percent.

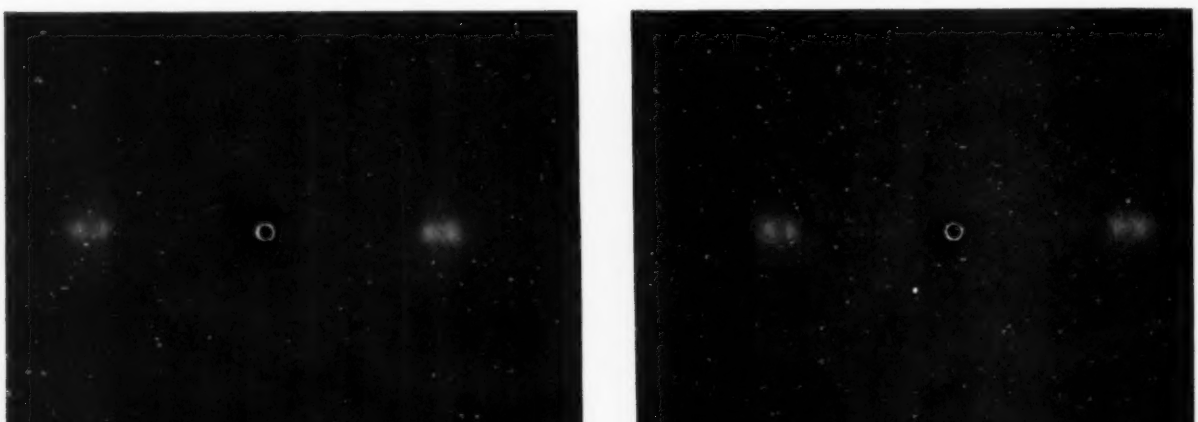
Photometer traces were taken of the spots, arcs and halo of the diffraction patterns. A full interpretation of these patterns involves tedious and laborious procedures. The present paper is concerned with their more direct and obvious features which are pertinent to a plastic deformation study. In all cases attention was confined to relative changes in the diffraction patterns which could be considered to result from plastic flowing.

RESULTS

With deformation it was observed that the diffraction pattern showed (1) changes in the dimensions of the fiber spots and (2) a decrease in the intensity of the amorphous relative to crystalline scattering.

Dimensional changes were observed in the two most prominent equatorial spots and these were of two types, which are designated as angular and radial broadening. The annular width at half-maximum intensity in a direction radially from the center of the pattern was taken as the radial broadness. It has been used by many investigators[3] for determining particle size or for detecting inhomogeneous lattice strains. One can show that the radial broadness of a diffraction spot may result from one or both of two sources: (1) Diffraction from small crystallographically ordered domains which is a function of the domain size, and (2) Diffraction from lattice planes with irregularly varying parameter possibly associated with the existence of internal strains which vary from point to point in the lattice. Simple theories of both types of radial broadness effects are essentially the same, and lead to relations for the dependence of the broadness on the Bragg angle and wave-length of the incident radiation which are identical to a multiplicative constant.[4] However, it will be seen that the data reported below lend themselves more readily to the particle size interpretation.

The angular broadness is defined as the width of the diffraction spot or line at half-maximum intensity taken in a direction at right angles to the radial broadness measurement; that is, the photometer trace is made around the pattern at a constant radius from the center. The angular broadness is taken as a measure of the degree of preferred orientation of the crystallites



a b

FIG. 2. Fiber patterns after relaxation at constant strain (*ca.* 26 percent). (a) Within 1 hour after termination of straining; (b) 48 hours later.

[3] S. D. Gehman and J. E. Field, J. App. Phys. **15**, 371 (1944).
[4] W. L. Bragg, Proc. Camb. Phil. Soc. **45**, 125 (1949).

in the fibers. If the orientation of the crystallites is broken down by heating or wetting of the fiber specimen, then the spots will diverge into more complete and uniform rings. As orientation or fibering takes place, the arcs converge into spots of smaller and smaller angular width. The interpretation of this broadness as a measure of the degree of orientation is clear and unambiguous.

Figure 3 shows a typical load-elongation curve for the Nylon used, when loaded at a constant time rate. After a small initial elastic rise followed by a somewhat rounded "yield point," the curve consists of two large portions; a concave upward region terminating in

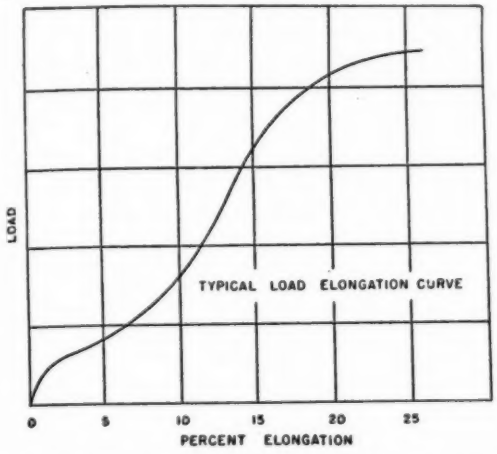


FIG. 3. Typical load elongation for Type 200 Nylon (constant rate of load application). Load units arbitrary.

an inflection point at about 14 percent elongation, and a following portion concave to the strain axis which terminates at the rupture point.

Figure 4 is a plot of the results of the angular broadness measurement; they are plotted against the same strain ordinate as in the previous diagram. The broadness decreases up to strains at the inflection point of the load-elongation curve, at which there appears to be a cusp, followed by an increasing broadness to the rupture point. From the interpretation of angular broadness cited above, this plot suggests, for strains up to the inflection point of the load-elongation curve, that the crystallites rotate toward positions parallel to the fiber axis. Further, these data suggest that this process terminates at the inflection point and is replaced by a disordering process. This interpretation is supported by the data shown in Fig. 5. Here the radial broadness is plotted as a function of the elongation. For strains up to the inflection point of the load-elongation curve, the broadness remains approximately constant but rises rapidly for larger strains. From the shape of this curve and from Fig. 3, it seems rather difficult to attribute this radial broadening to tessellated elastic strains, existing throughout the irradiated volume, which would be expected to rise monotonically throughout the straining epoch. On the other hand, it appears logical to associate this with a reduction in particle size, that is, crystallite break-up. Analogous measurements of radial broadening associated with macroscopic plastic torsional strains have been made for metals by Patterson and Orowan.[5] In both the magnitude of the broadness change and in shape, their results are similar to those broadness values shown in Fig. 5 which are associated with strains greater than the inflection point of the Nylon load-strain curve, 14 percent. The view of these authors varies as to the details of interpretation of this result.

The intensity of the amorphous scattering was compared to the crystalline scattering and was found to decrease with strain up to the inflection point of the load-elongation curve, and to remain approximately constant thereafter. Since this measure was difficult to make accurately, we offer at present only the qualitative result.

DISCUSSION

The experimental result and its apparent interpretation which were obtained above can be used together with other known observations concerning plastic flowing in metallic solids to infer a very interesting and probably general characteristic of the reaction of reasonably well-ordered solids to large load applications.

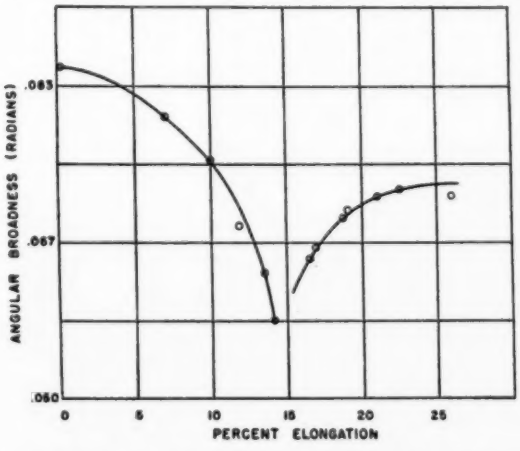


FIG. 4. Angular broadness vs. the elongation (%). Type 200 Nylon.

Recently, Koehler and Blewitt,[6] from studies of the changes in electrical resistivity with applied load, have inferred the existence of a sharp change in the micro deformation mechanisms at small plastic strains in ordered $AuCu_3$ single crystals. They have further pointed out that a similar inference can be obtained from: (1) a study of the curvature of "resolved" stress-strain relations of single crystals of $AuCu_3$, Cu, Ag, and Au, (2) x-ray line broadness changes with cold work in polycrystalline Cu, and (3) in other less direct ways. In view of the Nylon result reported above, the

[5] M. S. Patterson and E. Orowan, Nature **162**, 991 (1948).
[6] J. S. Koehler and T. H. Blewitt, Phys. Rev. **75**, 1952 (1949).

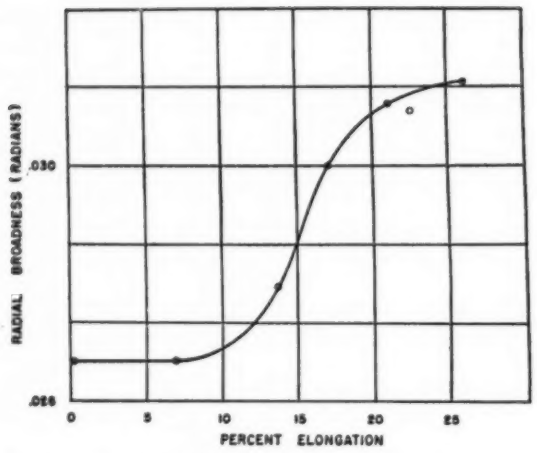


FIG. 5. Radial broadness *vs.* elongation (%). Type 200 Nylon.

Koehler-Blewitt generalization might be revised to the following. For any well-ordered solid, one will likely find in the region of true plastic flowing a narrow region of plastic strain through which there occurs a relatively sharp transition in the micro mechanisms of reaction to applied loads.

Recently, many theoretical papers have appeared which attempt to explain plastic flowing, and apparently related phenomena in metallic solids in terms of elementary units of crystallographic disorder termed "dislocations."[7,8] In view of the relatively high degree of order in the Nylon used for the work reported in this paper, one might seek an explanation of the results presented in terms of the dislocation concept. It should be noted, however, that the broadness measure reported here does not constitute evidence for the existence of particular dislocation arrays or, in fact, any array within a domain, but rather the existence of domains characterized by a linear dimension which is inversely proportional to the broadness. This dimension is the quantity which changes with the applied load. The domain viewpoint focuses, then, one's attention on the physically observed, whereas an explanation of the Nylon effects in terms of a generalized dislocation concept would stress entities which were not observed in the experiments. It is recognized that some type of disorder must exist between domains leading to x-ray line broadness, but it is felt that until experimental techniques are perfected which allow its study, explanations of general mechanical behavior in terms of disorder concepts should be avoided.

The suggestion that there is generally a transition in the microscopic mechanisms of plastic flowing in solids will, if verified, pose important theoretical problems. Currently, theoretical discussions of work hardening, e.g., in terms of dislocations often assume what is essentially an "exhaustion" hypothesis. Given amounts of the disorder are thought to initially exist. These theories explain the effects associated with continued plastic flowing in terms of the diminution of the number of elementary disordered regions with increased strain or time. Such exhaustion theories will surely require modifications if they are to account for the possibility of a transition in the micro reactions to gross flowing. The physical nature of flowing in solids, whatever it is, is surely extremely complex, even in what might be termed "simple" cases.

[7] G. I. Taylor, Proc. Roy. Soc. **A195**, 362 (1934).
[8] J. M. Burgers, Proc. Phys. Soc. **52**, 23 (1940).

CONCLUSIONS

The new data for highly oriented Nylon fibers presented indicate a change in the nature of plastic flowing, which occurs in the straining cycle at the point of inflection of the load-elongation curve (approximately 14 percent elongation). In the initial portion of the curve, the mechanism of flowing appears to involve rotation of the crystallites toward positions parallel to the fiber axis accompanied by an additional crystallization at the expense of the amorphous material initially present. The latter follows directly from the variations noted in the ratio of the intensity of the amorphous halo to that of the diffraction spots. This process terminates at the inflection point and is replaced by a mechanism of crystallite break-up, and disorientation.

The x-ray line broadness effects in Nylon, coupled with less direct effects observed in metallic solids which appear to be related to plastic straining all point to the existence in these solids of at least two types of microscopic reactions resulting from the load application. The experimental evidence, while not as direct as might be desired in some cases, seems to point to the existence of domains of nearly perfect atomic order, the motions and sizes of which are altered by the macroscopic flowing. Finally, from the domain viewpoint, it would appear that there are more general similarities between the micro reactions of well-ordered solids to applied loads than is commonly accepted. Obviously, domain structure on the atomic scale varies with materials, but it may be that this scale of structure is unimportant in describing at least some of the qualitative features of microscopic plastic flowing reactions.

ACKNOWLEDGMENTS

The authors are indebted to their colleagues G. R. Irwin and W. H. Sanders for many illuminating and helpful discussions. H. Friedman and L. S. Birks kindly made diffraction equipment in their laboratory available for this study.

Aluminum Oxide Replicas for Electron Microscopy Produced by a Two-Step Process

G. Hass and M. E. McFarland
Engineer Research and Development Laboratories, Fort Belvoir, Virginia
(Received October 26, 1949)

A two-step method is described for preparing aluminum oxide replicas of aluminum and non-aluminum specimens. In this method a thick layer of aluminum is evaporated onto the surface to be studied and pulled off the surface with Scotch Tape or cement covered glass. Aluminum oxide is formed on the side of the aluminum film which has been in contact with the specimen by anodizing in 3 percent ammonium tartrate at about 40 volts. The aluminum oxide film, which represents a positive replica of the specimen surface, is released from its substrate in dilute hydrochloric acid, washed with distilled water, and mounted for the electron microscope investigation. The applicability of the described method is demonstrated with electron micrographs of various specimens. The aluminum oxide films produced in ammonium tartrate are free of pores and grain structure. Films only 100 to 150A in thickness can be prepared and are suitable as temperature resistant supporting membranes for electron microscope and electron diffraction investigations.

ALUMINUM oxide coatings produced anodically in a suitable electrolyte have proved to be particularly useful as replica films for electron microscope studies.[1] Originally, however, the application of this method was restricted to aluminum and aluminum alloys on which anodic films can be formed directly. Only recently there was reported a modification of this method, a two-step process, applicable to studies of the surfaces of some non-aluminum materials.[2] In this process a piece of electrolytically polished pure aluminum is pressed against the surface to be studied; thus an aluminum replica is produced. On this aluminum replica a thin anodic film is produced and then removed chemically for examination in the electron microscope. Evidently this "engraving method" is only applicable to studies of extremely hard surfaces which are not damaged or changed by the high pressure required to press the aluminum against them.

By using a double replica method similar to the silver-collodion one first suggested by Zworykin and Ramberg,[3] a more general application of anodic aluminum oxide as replica film can be obtained. The specimen to be investigated is placed in a vacuum chamber and a thick film of aluminum ($t>1\mu$) and a thin layer of magnesium ($t<0.1\mu$) is evaporated onto its surface. The specimen is removed from the vacuum chamber and the magnesium-aluminum film is pulled off the surface with Scotch Tape or a cement covered glass slide. If the aluminum is sufficiently thick, the magnesium-aluminum layer can readily be removed from most specimens without damaging their surfaces. Aluminum oxide is now formed on the

FIG. 1. Unshadowed aluminum oxide replica of a 30,000 lines per inch grating. 20,000×.

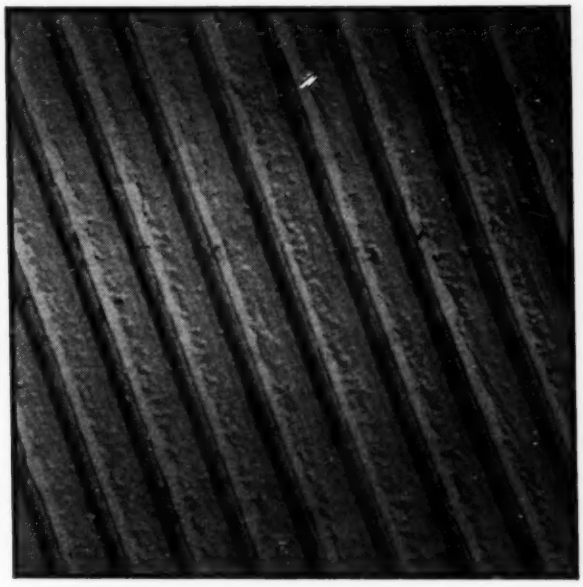


FIG. 2. Germanium-shadowed aluminum oxide replica of a 30,000 lines per inch grating. 10,000×.

[1] H. Mahl, Metallwirtschaft **19**, 488 (1940) and Zeits. Tech. Phys. **22**, 33 (1941).
[2] J. Hunger and R. Seeliger, Neue Phys. Blätter **2**, 15 (1946) and Metallforsch **2**, 65 (1947).
[3] V. K. Zworykin and E. G. Ramberg, J. App. Phys. **12**, 692 (1941).

FIG. 3. Germanium shadowed aluminum oxide replica of a Plexiglass surface scratched by cleaning with cotton. 21,000×.

FIG. 4. Germanium shadowed aluminum oxide replica of a KRS-5 crystal on which KRS-5 has been condensed. 14,000×.

side of the aluminum film which has been in contact with the specimen by anodizing with 30–40 volts for 20–30 sec. The electrolyte used for the anodic oxidation is a solution of 3 percent tartaric acid with ammonium hydroxide added to make the *p*H about 5.5. The thin magnesium film placed between the aluminum and the Scotch Tape or cement makes the separation of the aluminum oxide replica easier and faster than it would be with aluminum only. If the film now consisting of aluminum oxide, aluminum, and magnesium on Scotch Tape is immersed in dilute hydrochloric acid the magnesium dissolves rapidly and the aluminum film with the oxide replica on top rises to the surface in less than a minute. There the aluminum also dissolves in a few minutes and the aluminum oxide replica film is left floating on the surface. After being washed in distilled water, the replica film can be mounted on a support screen for the electron microscope investigation. Replicas produced by the above method can be shadow cast to increase the contrast of surface details. Germanium has been used for shadowing because it evaporates easily and forms films with glass-like structure when condensed onto unheated bases.[4,5] Figures 1 and 2 are unshadowed and germanium-shadowed aluminum oxide replicas of a 30,000 lines/inch grating obtained by the aluminum-aluminum oxide method. The replicas were made from a reflection grating ruled on an evaporated aluminum mirror. It is remarkable that a thick film of aluminum evaporated onto the aluminum grating can be removed without damaging the grating. Aluminum-aluminum oxide replicas have been made from the original grating more than 30 times without changing its appearance. The unshadowed grating replicas (Fig. 1) show sharp dark lines and are therefore especially suitable for calibrating the magnification of electron microscopes and for measuring particle sizes exactly. The fact that anodic aluminum oxide films are of uniform thickness and of temperature resistant material makes the replicas withstand electron bombardment of high intensity. The germanium shadowed replica (Fig. 2) shows more contrast and detail and allows a better study of the grating surface. Figures 3 and 4 demonstrate the general applicability of the two-step replica method described above. Figure 3 shows an electron micrograph of a Plexiglass surface scratched by careful cleaning with cotton and Fig. 4 represents the surface of a KRS-5 crystal (TlBr-TlI) on which KRS-5 has been condensed slowly at high temperature.

The thickness of the oxide films formed anodically in a given time on aluminum in 3 percent ammonium tartrate increases linearly with the applied voltage. Layers produced in 30 sec. have a thickness of 12.2A-units per volt. After 30 sec. their thickness increases very little with the duration of the anodic treatment.[6] Such layers are amorphous and show no evidence of pores or grain structure. Films only 100 to 150A in thickness can be prepared. They are, therefore, suitable as temperature resistant support films for electron diffraction and electron microscope investigations.[7] Supporting membranes with extremely smooth surfaces can be prepared on electrolytically polished pure aluminum or on thick evaporated aluminum layers that have been stripped from a glass surface.

[4] H. König, Reichber. f. Physik **1**, 4 (1944).
[5] G. Hass, Phys. Rev. **72**, 174 (1947).
[6] G. Hass, J. Opt. Soc. Am. **39**, 532 (1949).
[7] G. Hass and H. Kehler, Kolloid Zeits. **95**, 26 (1941).

Diffusional Viscosity of a Polycrystalline Solid

CONYERS HERRING
Bell Telephone Laboratories, Murray Hill, New Jersey
(Received December 23, 1949)

According to a suggestion of Nabarro, any crystal can change its shape by self-diffusion in such way as to yield to an applied shearing stress, and this can cause the macroscopic behavior of a polycrystalline solid to be like that of a viscous fluid. It is possible that this phenomenon is the predominant cause of creep at very high temperatures and very low stresses, though not under more usual conditions. The theory underlying it is developed quantitatively, and calculations of rate of creep, or equivalently of effective viscosity, are given for aggregates of quasi-spherical grains and for wires composed of cylindrical grains. Allowance is made for the effect of tangential stress relaxation at the grain boundaries. It is suggested that mosaic boundaries and boundaries between grains of nearly the same orientation may be unable to serve as sources or sinks of the diffusion currents, in which case the creep rate will depend only on the configuration of grain boundaries having a sizable orientation difference. Numerical comparison of the theory with experiments on the high temperature creep of wires favors this view, but is not entirely satisfactory. Suggestions for further experiments are made.

NABARRO[1] has recently pointed out that self-diffusion within the grains of a polycrystalline solid can cause the solid to yield to an applied shearing stress, the yielding being caused by a diffusional flow of matter within each crystal grain away from boundaries where there is a normal pressure and toward those where there is a normal tension, as shown in Fig. 1. This yielding is macroscopically describable by an effective viscosity proportional to the square of the linear dimensions of the grains. This phenomenon, which occurred independently to the present writer, provides a possible explanation for the quasi-viscous behavior which has been observed for foils and wires which are suspended with a very slight load and held at an elevated temperature.[2] It may also be important in other types of sintering phenomena. The purpose of the present note is to clarify certain points in the theory underlying this "diffusional viscosity" and to attempt a more refined quantitative calculation of its magnitude for several experimentally realizable situations.

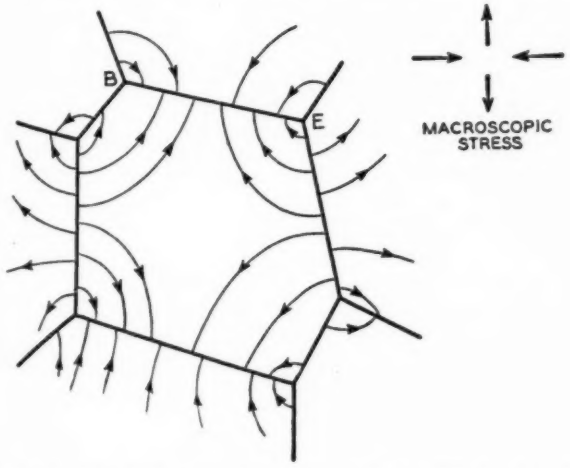


FIG. 1. A typical grain of a polycrystalline solid, showing the self-diffusion currents to be expected when the solid is subjected to a shearing stress.

1. BASIC THEORY

Let us consider first the simplest case of stress-motivated diffusion, namely self-diffusion in a single crystal of the cubic class containing only one type of atom and held at uniform temperature. In the absence of pressure gradients the diffusive flux of atoms, due to migration of lattice vacancies or interstitial atoms, will be proportional to the gradient of the concentration of these lattice defects; in the presence of pressure gradients, however, it tends to be energetically advantageous to move the lattice defects in whichever direction will relieve the inequality of pressure. It is easily shown that the effects both of concentration gradients and of pressure gradients are correctly given by writing the diffusive flux $\mathbf{j}$ (atoms per unit area per unit time) as proportional to the gradient of a chemical potential.[3,4] If n_L is the number of lattice sites per unit volume, D the self-diffusion coefficient, and T the absolute temperature,

$$\mathbf{j}=-(n_L D/kT)\nabla(\mu-\mu_h), \qquad (1)$$

where $(\mu-\mu_h)$ is the work per atom required for the reversible addition of a small quantity of material to the given region of the crystal, the material added being placed interstitially in the lattice or used to fill lattice

[1] F. R. N. Nabarro, "Report of a Conference on the Strength of Solids" (Physical Society, London, 1948), p. 75.

[2] *Foils:* J. C. Chapman and H. L. Porter, Proc. Roy. Soc. **A83**, 65 (1909); I. Sawai and M. Nishida, Zeits. f. Anorg. u. Allg. Chemie. **190**, 375 (1930); G. Tammann and W. Boehme, Ann. d. Physik **12**, 820 (1932). *Wires:* Udin, Shaler, and Wulff, J. Met. **1**, 186 (1949); Alexander, Kuczynski, and Dawson, paper presented at the Symposium on the Physics of Powder Metallurgy organized by Sylvania Electric Products, Inc., New York, August 24–26, 1949.

[3] It has long been recognized that it is more appropriate to write the diffusion equations for multi-component systems with chemical potentials, or their equivalents, rather than with concentrations, and this fact has recently been the subject of a number of papers in the metallurgical literature. The most recent, and probably the most satisfactory, of these is that of J. Bardeen, Phys. Rev. **76**, 1403 (1949).

[4] C. Herring, paper presented at the Symposium on the Physics of Powder Metallurgy organized by Sylvania Electric Products, Inc., New York, August 24–26, 1949. Plans have been made to publish the papers of this Symposium in book form.

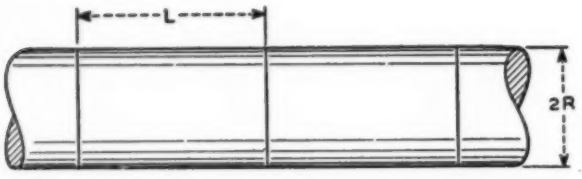


FIG. 2. Section of a wire with equally spaced transverse grain boundaries.

vacancies. This work is represented as a difference of two chemical potentials, μ and μ_h, to emphasize the fact that the change in question alters both the number of atoms and the number of lattice defects; however, in this paper only the difference of these quantities will appear.

Since any divergence of $\mathbf{j}$ within a single crystal would rapidly alter the concentration of lattice defects in such direction as to reduce $|\nabla\cdot\mathbf{j}|$, it will be safe to assume that in any continuing change $\nabla\cdot\mathbf{j}\propto\nabla^2(\mu-\mu_h)=0$. Thus the diffusion currents are completely determined if the values of $(\mu-\mu_h)$ on the boundary of the crystal are known. Now if the transfer of atoms from the interior to lattice sites at the boundary can occur easily, i.e., if the surface of the crystal is free to grow or recede with an accompanying change in the number of lattice defects in the interior, the boundary values of $(\mu-\mu_h)$ can be determined from the condition that the free energy of the crystal and its surroundings be a minimum with respect to transfer of atoms between any portion of the surface and the volume immediately beneath this portion. In reference 4 the boundary values of $(\mu-\mu_h)$ are determined in this way for the free surface of a crystal, and for grain boundaries where the interfacial tension between the grains is important; here we shall consider the simpler case where surface free energies can be neglected in comparison with the effects of externally applied forces. In intermediate cases the two effects have merely to be added.

Assuming that the surface-volume equilibrium just mentioned is realized at a grain boundary, the work involved in adding δN atoms to a crystal grain in the form of a hump on its surface must turn out to be the same, to the first order in δN, as the work involved in putting δN atoms into the interior just below the same surface region, holding the position of the surface fixed and placing the new atoms in interstitial positions or in places formerly occupied by lattice vacancies. The latter work is $(\mu-\mu_h-\mu_0)\delta N$, where μ_0 is the chemical potential of the source from which the δN atoms are obtained, say a piece of the same substance in equilibrium at the same temperature as the crystal being considered and at zero stress. The other process, formation of a surface hump, involves an amount of work $(\delta F-\mu_0\delta N-p_{zz}\delta v)$, where δF is the free energy of the hump itself, δv its volume, and p_{zz} is the normal traction on the surface, i.e., the zz component of the stress-tensor, where the z axis is chosen normal to the surface. Now to the first order in the stresses, $\delta F=\mu_0\delta N$, so equating the two amounts of work gives

$$\mu-\mu_h=\mu_0-p_{zz}\Omega_0, \tag{2}$$

where $\Omega_0=\delta v/\delta N$ is the atomic volume. This is the generalization of Nabarro's result regarding the effect of normal stress on vacancy concentration at the surface.

One can easily generalize Eqs. (1) and (2) to non-cubic crystals or to crystals containing more than one kind of atom, provided the composition is uniform. In the multi-component case, however, care must be exercised in attempting to relate the D occurring in (1) to diffusion coefficients measured in other ways. Although a detailed discussion of diffusion in multi-component systems would be beyond the scope of this paper, an attempt has been made in the Appendix to indicate the type of analysis which must be used in applying the present equations to such things as substitutional solid solutions.

2. SPECIFIC CALCULATIONS

The rate of yielding of any given specimen to applied forces obviously depends on the detailed distribution of the sizes and shapes of its crystal grains, and on whether or not the grain boundaries are able to sustain shearing stresses for times as long as are required for appreciable diffusional transport to take place. It has been established in a number of cases that shearing stresses across metallic grain boundaries are rapidly relaxed at high temperatures;[5,6] it seems likely that this relaxation is a general property of grain boundaries, and that until this relaxation has become complete, the rate of creep due to it will usually be much faster than that due to diffusional viscosity. Although it is thus much more reasonable to compute the diffusional viscosity on the assumption of relaxed tangential stresses, we shall give calculations for both relaxed and unrelaxed tangential stresses, to show how much difference this factor makes. Both these calculations will be made assuming uniform, roughly spherical grains of a size small compared with that of the specimen; to show the way in which the rate of yielding is affected by having grains comparable with the dimensions of the specimen, some calculations will then be given for the rate of elongation of a wire consisting of roughly cylindrical grains end to end. Most of these cases are of course too idealized to apply accurately to any actual case, but it is hoped that they will sometimes be fair approximations to the truth.

Equal Quasi-Spherical Grains, No Grain Boundary Flow

If a crystal made up of grains such as that of Fig. 1 were suddenly subjected to a shearing stress, the initial diffusion currents would be very large in the neighborhoods of all sharp edges, such as E in the figure. This means that any grain boundary such as BE which is

[5] T. S. Kê, Phys. Rev. **71**, 533 (1947); J. App. Phys. **19**, 285 (1948); **20**, 274 (1949).
[6] King, Cahn, and Chalmers, Nature **161**, 682 (1948).

subjected to a normal tension will have more extra matter squeezed in at its edges than at its center. This will increase the tension at the center and decrease that at the edges, and a situation will very soon be reached where the normal flux is fairly uniform over each facet of the boundary. This suggests that for a crystal made up of nearly spherical grains all of the same size and having nearly isotropic elastic properties the rate of yielding can be fairly well approximated by that of a sphere acted on by such surface tractions as will produce a uniform shearing stress in its interior.

We shall assume for simplicity that the equations of Section 1 for a cubic crystal apply. Let the macroscopic stress-tensor be $\bar{p}_{\mu\nu}$. This must in all cases equal the volume average of the microscopic stress; in the present case we assume the stress in the interior of the sphere to be constant and equal to $\bar{p}_{\mu\nu}$. We assume $\bar{p}_{\mu\nu}$ to be a pure shear, so that

$$\sum_{\mu} \bar{p}_{\mu\mu}=0.$$

At a point of the surface of the sphere with coordinates x_μ the normal traction is

$$p_{rr}=\sum_{\mu,\nu} \bar{p}_{\mu\nu}x_\mu x_\nu/R^2,$$

where R is the radius of the spherical grain, so that

$$\sum_{\mu} x_\mu^2=R^2.$$

Inserting this into the boundary condition (2) it is easily seen that since $\nabla^2(\mu-\mu_h)=0$ the chemical potential in the interior must be given by

$$\mu-\mu_h=\mu_0-\Omega_0\sum_{\mu,\nu} \bar{p}_{\mu\nu}x_\mu x_\nu/R^2. \tag{3}$$

The normal flux at the boundary is, by (1)

$$j_n=(dr/dt)/\Omega_0=(2D/kT)\sum_{\mu,\nu} \bar{p}_{\mu\nu}x_\mu x_\nu/R^3 \tag{4}$$

since $n_L\Omega_0=1$. This may be compared with the rate of normal displacement of the boundary in a uniform shear of strain rate $\dot{u}_{\mu\nu}$:

$$dr/dt=\sum_{\mu,\nu} \dot{u}_{\mu\nu}x_\mu x_\nu/R. \tag{5}$$

Comparison of (4) and (5) shows that our assumptions are consistent with a uniform shear of all the grains of the crystal at a rate

$$\dot{u}_{\mu\nu}=(2D\Omega_0/kTR^2)\bar{p}_{\mu\nu}. \tag{6}$$

Viscous flow is defined, for the present case where $\bar{p}_{\mu\nu}$ is a pure shear, by the macroscopic relation $2\dot{u}_{\mu\nu}=\bar{p}_{\mu\nu}/\eta$, which with (6) gives for the effective viscosity of the polycrystalline solid

$$\eta=kTR^2/4D\Omega_0=(3/32\pi)^{\frac{1}{3}}(kT/D\Omega_0)V_g^{\frac{1}{3}}, \tag{7}$$

where V_g is the volume of a grain.

Equal Quasi-Spherical Grains, Tangential Stress Relaxed at Boundaries

For this case we must find a distribution of normal tractions over the surface of a sphere which will make the volume average of the stress-tensor in the interior equal the macroscopic stress-tensor $\bar{p}_{\mu\nu}$. Now from the equilibrium condition that the divergence of the stress-tensor must vanish we have, if $d\tau$ runs over the volume of the sphere and dS over its surface

$$0=\int x_\mu \sum_\alpha \partial p_{\alpha\nu}/\partial x_\alpha d\tau$$

$$=-\int p_{\mu\nu}d\tau+\int x_\mu \sum_\alpha p_{\alpha\nu}n_\alpha dS, \tag{8}$$

where $\mathbf{n}$ is the unit outward normal. For the present case

$$\sum_\alpha p_{\alpha\nu}n_\alpha=p_{nn}n_\nu,$$

so this gives

$$\bar{p}_{\mu\nu}=(3/4\pi R^4)\int x_\mu x_\nu p_{nn}dS. \tag{9}$$

If p_{nn} is represented by a series of spherical harmonics, (9) will determine the coefficients of the harmonics of order 2 (the lowest order consistent with the conditions of the problem) but will leave the higher order terms arbitrary. In an actual crystal grain these higher order terms will probably be no larger than is required to take account of the departure of the grain boundary configuration from spherical symmetry, and it is therefore proper to neglect them in the present approximation, the more so as they make no contribution to the over-all shear rate. Let us therefore set $p_{nn}=A(x_1^2-x_2^2)/R^2$, a form which can always be realized for the contribution of the second order spherical harmonics by proper choice of the coordinate directions x_1 and x_2. Insertion of this into (9) gives $\bar{p}_{11}=-\bar{p}_{22}=2A/5$, all other $\bar{p}_{\mu\nu}=0$. The chemical potential in the interior of the sphere is given by

$$\mu-\mu_h=\mu_0-(5/2)\bar{p}_{zz}\Omega_0(x^2-y^2)/R^2. \tag{10}$$

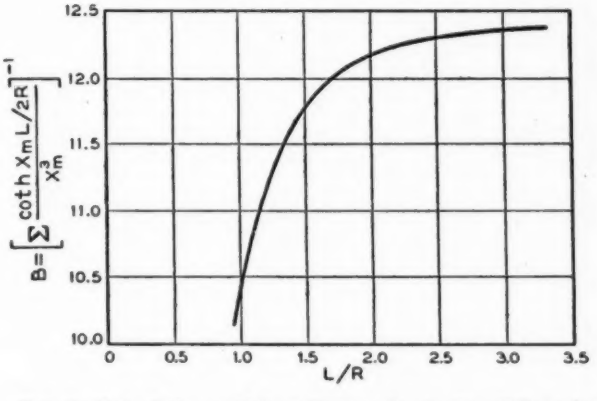


FIG. 3. Plot of the coefficient B occurring in Eq. (18) against the ratio of the length L of the grains to the radius R of the wire.

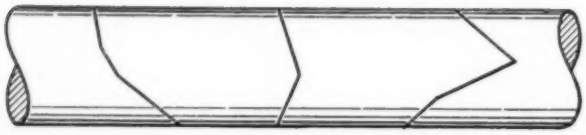


FIG. 4. Wire with irregular grain boundaries.

This is of the same form as (3) except for the factor 5/2. Repeating the argument leading to (7) gives

$$\eta = kTR^2/\tfrac{1}{10}D\Omega_0 = \tfrac{1}{10}(3/4\pi)^{\frac{1}{3}}(kT/D\Omega_0)V_a^{\frac{1}{3}}, \quad (11)$$

which is smaller than (7) by a factor $\frac{2}{5}$.

Wire with "Bamboo-Like" Grain Structure

Fine metal wires often consist of crystal grains occupying the entire cross section of the wire and having lengths which may be of the same order as the diameter of the wire or may be many times larger. In the experiments cited in reference 2 this type of grain structure characterized both the copper wires used by Udin, Shaler, and Wulff and the gold wires used by Alexander, Kuczynski, and Dawson.* A model which approximates the condition of these wires and for which the associated potential theory problem is easily solved is that of a wire with a "bamboo-like" grain structure, i.e., a cylinder of radius R with plane grain boundaries normal to the axis of the cylinder, uniformly spaced a distance L apart, as shown in Fig. 2.

For this model it is obvious from symmetry that when a longitudinal tension or compression is applied, the diffusion currents in the quasi-steady state will be such as to make the normal flux away from or into each grain boundary a constant over its area. This will of course require that the normal traction p_{nn} across a grain boundary vary with distance from the center of the wire; however, the integral condition

$$\int p_{nn}dS = F$$

must always be obeyed, where dS runs over the area of a grain boundary, and F is the force applied at the ends of the wire. For simplicity we ignore surface tension effects for the present; this is legitimate in computing the change in rate of stretching with a change in the load. In such case the boundary conditions on $(\mu-\mu_A)$ are:

$$f \equiv \mu - \mu_A - \mu_0 = 0 \quad \text{on the cylindrical surface} \quad (12)$$

while over the end boundary of each grain

$$\partial f/\partial z = \text{constant} \quad (13)$$

$$\int f dS = -F\Omega_0 \quad (14)$$

where z is the coordinate in the direction of the axis of the cylinder. The determination of f involves the same problem, therefore, as the determination of the potential due to an infinite series of uniformly charged disks placed inside a conducting cylinder. Using standard methods[7] the solution is easily shown to be

$$f = -\frac{F\Omega_0}{2\pi R^2} \frac{\sum_m [\coth(X_mL/2R)\cosh(X_mz/R) - \sinh(X_mz/R)]J_0(X_mr/R)/X_m^2J_1(X_m)}{\sum_m \coth(X_mL/2R)/X_m^3} \quad (15)$$

where the J's are Bessel functions, X_m runs over the roots of $J_0(X_m)=0$, and r is the distance from the axis of the cylinder, and where z ranges from 0 on one of the grain boundaries to $L/2$ midway between this boundary and the next. Each grain elongates or contracts at a rate

$$dL/dt = -(2\Omega_0/\pi R^2)\int j_z dS,$$

the integration being over the disk of radius R in the plane $z=0$. Using (1) and (15) we find

$$dL/dt = (D\Omega_0/2\pi R^3 kT)BF, \quad (16)$$

where

$$B = [\sum_m \coth(X_mL/2R)/X_m^3]^{-1} \quad (17)$$

is a function of L/R plotted in Fig. 3. Its asymptotic value for $L \gg R$ is $B_\infty = 12.37$, and it does not begin to depart seriously from this value until L/R has become smaller than any value likely to be encountered frequently in wires whose crystal grains occupy the whole cross section.

In the present case the deformation of the wire cannot legitimately be described as viscous, because the grain size equals or exceeds the diameter of the specimen; however, it is interesting for comparative purposes to compute the viscosity η which a viscous filament of the same diameter would have to have in order to elongate at the same rate under the same load. For such a filament we have $F = \pi R^2 \cdot 3\eta d\ln L/dt$, and this agrees with (16) if

$$\eta = 2kTRL/3BD\Omega_0. \quad (18)$$

Wire with Irregular or Slanting Grain Boundaries, Tangential Stresses Relaxed

It is easy to show that if the grain boundaries are not normal to the axis of the wire as in Fig. 2, but are of the type shown in Fig. 4, the rate of yielding of the wire will

* I am indebted to the authors of these two papers for discussions of unpublished details of their work, including opportunities to examine photomicrographs of their wires, and for the statistics on mean grain lengths which have been used in Table I.

[7] See for example, W. R. Smythe, *Static and Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1939), Chapter 5.

be greater than that given by (16) for the same grain size, if the tangential stresses vanish at the grain boundaries. This is most easily seen by noting that a given dL/dt requires a given total normal flux integrated over the grain boundary,† regardless of its shape or inclination; in the electrostatic analogue this means a given total charge. The potential produced by this given charge, averaged over the grain boundary, will be numerically smaller when the grain boundary is oblique than when it runs normally across the wire as in Fig. 2. This potential is proportional to the normal stress across the grain boundary. It is easy to see that when the normal tractions are resolved along the axis of the wire, equal mean normal stresses will give equal values of the force F tending to elongate or compress the wire; therefore F will be less for Fig. 4 than for Fig. 2, when dL/dt is the same.

The calculations of the preceding section suggest some approximations which can be used to make a crude quantitative estimate of this effect. The same methods which were used to obtain (15) can be applied to calculate the potential due to a uniformly charged flat ring of any inner radius, r_1, and outer radius r_2, placed inside a grounded conducting cylinder. The potential problem for a grain boundary consisting of a series of ring-like steps, as shown in Fig. 5, can therefore be solved, and if the steps are not taken too large, this solution can probably be regarded as typical of irregular grain boundaries of the same average inclination. Ignoring interactions between the two ends of a grain, which Fig. 3 shows to be quite small, the force F required to produce a given elongation dL/dt at a single grain boundary of the sort described comes out to be

$$F=\frac{2\pi R^3 kT}{D\Omega_0}\times\left[\sum_m \frac{\sum_{i,j}\{m,i\}\{m,j\}\exp(-|z_i-z_j|X_m/R)}{X_m^5[J_1(X_m)]^2}\right]\frac{dL}{dt}, \quad (19)$$

where i and j label the various rings, the ith ring having inner radius r_{i1} and outer radius r_{i2} and longitudinal coordinate z_i, and where

$$\{m,i\}=(X_m r_{i2}/R)J_1(X_m r_{i2}/R)-(X_m r_{i1}/R)J_1(X_m r_{i1}/R).$$

Now for large L the series in (17) differs from its leading term by only about 10 percent, so it is probably legitimate, in an approximation which replaces the grain boundaries of Fig. 4 by those of Fig. 5, to set the ratio of (19) to the F_0 given by (16) equal to the ratio of the first terms of the respective summations on m. This gives

$$F/F_0\approx\sum_{i,j}\{1,i\}\{1,j\}/X_1^2[J_1(X_1)]^2. \quad (20)$$

Some values computed from this formula for typical grain boundaries are given in Fig. 5.

The conclusion from this analysis is that the equivalent viscosity of a wire with irregular grain boundaries may be as much as two to three times smaller than the value given by (18) for flat-ended grains of the same length.

3. GRAIN BOUNDARIES *VERSUS* MOSAIC BOUNDARIES

The calculations of the preceding section have assumed the existence of certain "grains," at whose boundaries diffusion currents of lattice defects can converge or diverge, the boundary concentration of lattice defects being in equilibrium with the normal stress. It is conceivable that these properties might be possessed by crystal grains in the usual sense of the term, by mosaic blocks within a "single crystal," or, perhaps, by neither. In the present section it will be suggested as particularly likely that the condition of equilibrium just mentioned is attained, at sintering temperatures, on most ordinary grain boundaries, but is not attained on the boundaries of mosaic blocks. The experimental evidence discussed in the next section, though not entirely satisfactory, tends to favor this suggestion.

The boundaries of mosaic blocks have been pictured as consisting of rows of identical parallel dislocations spaced apart by a distance of the order of 10^3 atomic spacings.[8] Although it is not known whether the usual mosaic structure of single crystals consists of blocks bounded by such rows of dislocations, or merely of an irregular tangle of dislocation lines, it can hardly be doubted that the structure can be described in terms of dislocations; moreover, there seem to be good prospects for accounting quantitatively for many of the properties of grain boundaries, at least when the orientation difference between the grains is small, by picturing the boundaries as arrays of dislocations.[9] It is therefore

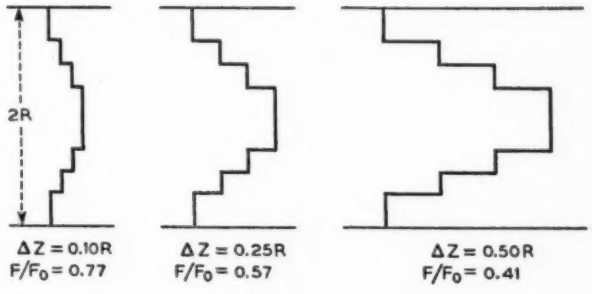


FIG. 5. Representation of irregular grain boundaries by successions of ring-like steps of height Δz at radii $0.3R$, $0.5R$, and $0.7R$. The quantity F/F_0 shown for each case is the approximate ratio of the force required to produce a given rate of elongation of the grain shown to the force required to produce the same rate of elongation of one of the flat-ended grains of Fig. 2.

† In accordance with the discussion given in Section 4, it is assumed that for sufficiently small stresses no "offsetting" of the grains will occur.

[8] J. M. Burgers, Proc. Phys. Soc. **52**, 23 (1940); J. S. Koehler and F. Seitz, J. App. Mech. **14**, A217 (1947).

[9] W. Shockley and W. T. Read, Phys. Rev. **75**, 692 (1949); W. T. Read and W. Shockley, Phys. Rev. **78**, 275 (1950).

legitimate to wonder what difference, if any, there is between the behavior of grain boundaries and that of mosaic boundaries.

In a well-annealed crystal the configurations of the dislocation arrays describing the mosaic structure must be such as to make the elastic free energy a minimum with respect to small displacements of any of the dislocation lines. Application of a small stress will affect the dislocations in two ways: a shearing stress on the slip plane of a dislocation will tend to move it in this plane, while a normal stress across a plane perpendicular to the slip vector will tend to move the dislocation in a direction not lying in the slip plane. Motion of the latter type requires the creation or annihilation of lattice defects, and will cease when the chemical potential of the lattice defects in the neighborhood of the dislocation has changed by an amount sufficient to compensate for the change in stress. An isolated dislocation line would thus affect the potential $(\mu-\mu_h)$ of the diffusion problem in much the same way as a wire held at a certain potential would affect an electrostatic problem. However, in an array of reasonably well-separated dislocations it seems likely that the freedom of motion of each dislocation will be impeded by the stress-fields of neighboring dislocations, and that within a short time after application of a small stress a new elastic equilibrium will be reached and all motion of the dislocations will cease, and concomitantly $(\mu-\mu_h)$ will become essentially constant over mosaic dimensions.

One aspect of this speculation can be given a rough check. The distance which each dislocation will move, in coming to a new position of equilibrium after application of a small stress σ, may be expected to be of the order of magnitude of $\sigma d/\sigma_0(d)$, where d is a length of the order of the spacing between dislocations and $\sigma_0(d)$ is the stress at a distance d from an isolated dislocation line. The resultant shearing strain of the whole specimen should therefore be of the order of this quantity times a/d^2, where a is the lattice constant. Since $\sigma_0(d)$ is of the order of $Ga/2\pi(1-p)d$, where G is a shear constant and p a Poisson's ratio,[10] the resulting shear is of order $2\pi(1-p)\sigma/G$. This gives values two or three orders of magnitude smaller than the total strains observed in the wire experiments of reference 2, which were of the order of 10^{-2}.

It is therefore not implausible to suppose that motion of dislocations may be of negligible importance in these experiments, and that the diffusion currents are divergenceless except at grain boundaries separating two crystals having a sizable difference of orientation $\Delta\theta$. By "sizable" is meant an orientation difference large enough so that the grain boundary can no longer be viewed as an array of dislocations spaced many lattice constants apart. If this speculation should turn out to be true, it would imply the existence of a critical value of $\Delta\theta$ below which growth of the two crystals normal to the grain boundary cannot take place freely, but above which it can. In such case any grain boundaries separating grains with less than the minimum $\Delta\theta$ should not be counted as grain boundaries in comparing the present theory with observations. There may conceivably be many such boundaries in wires whose grains have a strong tendency to crystallize with a certain crystallographic direction along the axis of the wire. For two grains with such a common axis need only a relative rotation about the wire axis to bring their orientations into coincidence, and it is possible that at sintering temperatures the natural desire of the grain boundary to decrease its interfacial tension may bring about such a rotation, at least until the orientation difference becomes small enough so that further changes of configuration cannot take place freely.

[10] See for example, J. S. Koehler, Phys. Rev. **60**, 397 (1941).

4. RELATION OF THEORY TO EXPERIMENT

Most of the present section will be devoted to a discussion of experiments on the longitudinal creep of wires at high temperatures, as measured by Udin, Shaler, and Wulff[2] for copper and by Alexander, Kuczynski, and Dawson[2] for gold. In these experiments an attempt was made to find what value of the applied tension would just suffice to prevent the wires from shrinking under the action of their surface tension, and thus to measure the latter quantity. The stresses used in these experiments were therefore below the range used in most ordinary creep experiments (about 1 to 100 g/mm^2), while the extensions were relatively large, of the order of 10^{-3} to 10^{-2} and more. It is therefore not unreasonable to attribute the creep in these high temperature experiments to mechanisms quite different from those operative in most of the previous creep experiments in the literature, although it is of course quite conceivable that the mechanisms may be the same, the thresholds being lower and the rates faster at the higher temperatures. The latter possibility is supported by the fact that Miller[11] has observed an apparently unlimited creep for single crystals of zinc under stresses of only 2 or 3 g/mm^2, even at room temperature; however, it would be risky to conclude from this that the experiments of reference 2 involve the same mechanisms as room-temperature creep in single crystals, and the following discussion will explore the alternative that the diffusion mechanism discussed in the previous sections plays a major role. As Nabarro has shown,[1] the latter mechanism does not seem to be the major one in creep experiments such as those of Chalmers[12] on single crystals of tin.

Offsetting and Kinking

Besides the diffusion effect and the possible effect of ordinary creep within the individual crystal grains, there are at least two other effects which may influence the changes in length which are observed for fine wires

[11] R. F. Miller, Trans. A.I.M.E. **122**, 176 (1936).
[12] B. Chalmers, Proc. Roy. Soc. **A156**, 427 (1936).

suspended at high temperatures. One of these is "offsetting" of adjacent grains, as shown in Fig. 6. This is known to occur for polycrystalline wires suspended at high temperatures under tension of sufficient magnitude.[13] However, it is easy to see that it cannot occur if the tension is below a certain limit, and that it cannot contribute to the shrinking of wires suspended with no load. For, referring to Fig. 6, consider a lateral offset of amount AB along a grain boundary whose normal makes an angle θ with the axis of the wire. This enables the wire to yield to the applied tensile force F by a distance $\langle AB\rangle \sin\theta$; however, it exposes an area of new surface represented by AB and CD. The surface free energy thus increases by $2R\cdot\langle AB\rangle\cdot(2\bar{\gamma}-\gamma_b)$ where $\bar{\gamma}$ is an average surface tension for the regions AB and CD, and γ_b is the specific free energy of the grain boundary. Thus the over-all free energy will be lowered by the offset if and only if

$$F>2R(2\bar{\gamma}-\gamma_b)\csc\theta. \tag{21}$$

This is somewhat larger than the force $\pi R\bar{\gamma}'$ necessary to prevent shortening of a long grain under its own surface tension. We may therefore conclude that offsetting will not occur in the lower range of stresses employed in the wire experiments of reference 2, but that it may well occur in the higher part of the stress range.

The other possible perturbing effect, which will be called "kinking," consists in the development of a slight angle between the axes of adjacent grains of the wire. This could be produced by a relative rotation about an axis normal to the wire axis. Such a rotation might be motivated by a dependence of the free energy of the grain boundary on the relative orientations of the grains; it would of course require that material be removed from the portions of the grains which move toward each other, and added to the portions which move apart. If this transport of matter takes place by diffusion, one can estimate the order of magnitude of the rate of change of orientation by the methods described in reference 4. This estimate, the details of which will not be given here, shows that for the wires of reference 2 the rate of kinking due to this cause ought to be too small to produce a perceptible shortening, though possibly not too small to give a measurable angle of kink.

Kinking might conceivably also be caused by a relative rotation of two grains about the normal to a plane grain boundary, if this normal were a trifle inclined to the axis of the wire. The motivation could again come from an anisotropy of the grain boundary energy; for this process, however, no transport of matter would be required, and its rate might be faster than that of the process described in the preceding paragraph.

Calculated *versus* Observed Creep Rates

Table I gives a comparison of the observed rates of creep of the wires of reference 2, expressed in terms of equivalent viscosities, with rates calculated from (18) and (20). Since the least accurately known of the experimental data are the lengths L of the crystal grains, the comparisons are made on this quantity.

Ideally, if the diffusion hypothesis is the correct one, it should be possible to get a fairly accurate check between theory and experiment. For, with one possible exception, all the quantities on the right of (18) can be measured fairly accurately, the most uncertain being D, and even this can be measured to an accuracy of 20 percent or better if the temperature can be accurately reproduced. The possible exception is the mean length L of the grains: if, following the possibility suggested in Section 3, it turns out that growth can occur at grain boundaries where the orientation difference is large but not at those where it is small, then the value of L to be inserted into (18) will be larger than the value L' measured metallographically, by an amount which it may be hard to establish. The correction factor F/F_0 will in many cases be close to unity, since both theory and observation agree that grain boundaries usually prefer to align themselves normal to the wire axis; if for any reason the boundaries are not transverse, a reasonable estimate of F/F_0 can be made from Fig. 5.

A number of experimental uncertainties in the present examples make the accuracy of the comparison of theory and observation fall somewhat short of this ideal. Besides the uncertain possibility that L may be appreciably greater than L', the value of L' itself is somewhat uncertain. The values given in the table were taken from photomicrographs of short sections of the wires after etching; however, since regions of long grains may alternate with regions of short grains, many such samples must be measured to get a reliable measure of the mean length of all the grains in the specimen. Moreover, the grain structure may change appreciably during the experiment. In this connection it is interesting to speculate that a relative axial rotation of neighboring grains, resulting ultimately in making the orientation difference too small for growth to occur, might conceivably cause the effective L to increase with time. This would lead to a creep rate decreasing with time, as was observed by Alexander, Kuczynski, and Dawson.[2]

A moderate error in the temperature, or rather, in the relative temperature scales of the creep and diffusion measurements, can produce an appreciable error in D. Dr. Alexander has informed the author that the temperature listed in the table as 1050° may have been closer to 1025°; a temperature lowering of this amount would

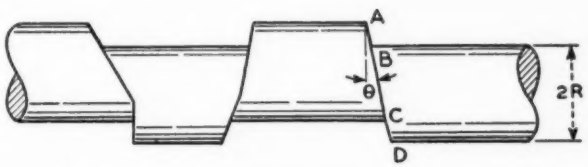


FIG. 6. Illustration of offsetting for a wire suspended under tension.

[13] See for example, C. J. Smithells, *Tungsten* (D. Van Nostrand Company, Inc., New York, 1936), second edition, p. 81.

TABLE I. Comparison of observed viscosities of wires with predictions of Eqs. (18) and (20).*

Authors†	USW		AKD
Metal	Cu	Cu	Au
T°C	1000	1000	1050
R, microns	36	64	14
D, cm²/sec.‡	2.5×10^{-9}	2.5×10^{-9}	8×10^{-9}
Observed η, poises	4.4×10^{12}	3.2×10^{12}	1.4×10^{12} (initial) 2.4×10^{12} (final)
$(L/R)(F/F_0)$, calculated	1.0	0.24	9.6 to 16
L'/R, observed	4	2	2

* The observed η-values for copper were taken from lines drawn by Udin, Shaler, and Wulff averaging their measurements at different temperatures. Their values must be multiplied by 2 because of an error in a formula, derived by Frenkel, by which they reduced their observations.

L is the average spacing between grain boundaries at which growth can occur, while L' is the average spacing between observable boundaries.

F/F_0 is the correction factor to be applied to (18) due to departure of the grain boundary shapes from the ideal transverse planes of Fig. 2; according to (20) this factor is probably between $\frac{1}{2}$ and 1.

The next to the last line of the table was computed by equating the observed viscosity η to the product of F/F_0 by the right of (18) taking B to be 12.2, and solving for $(L/R)(F/F_0)$.

† See reference 2.

‡ The self-diffusion coefficient tabulated for copper is close to the values measured by Steigman, Shockley, and Nix, Phys. Rev. **56**, 13 (1939); M. S. Maier and H. R. Nelson, Trans. A.I.M.E. **147**, 39 (1942); and Raynor, Thomassen, and Rouse, Trans. Am. Soc. Metals **30**, 313 (1942). That for gold was obtained by extrapolating the data of H. A. C. McKay, Trans. Faraday Soc. **34**, 845 (1938).

lower D, and therefore the calculated $(L/R)(F/F_0)$, by a factor of about 1.5.

If the diffusion of a radioactive tracer isotope occurs largely or predominantly by simultaneous interchange of positions by neighboring atoms, the D to be used in the present calculations will of course be less than that measured in the tracer experiment, perhaps very much less.

Remarks on Future Experiments

The comparisons given in Table I, though not very satisfactory, do at least show that the creep observed in the experiments of reference 2 is of the same order of magnitude as that predicted on the hypothesis that it is due to diffusion currents flowing between the free surface and the grain boundaries and that it is therefore distinctly slower than that which would be predicted by an analogous theory in which the diffusion currents converge on the boundaries of mosaic blocks with dimensions of the order of 10^{-4} cm. However, an obvious direct check of the hypothesis that grain boundaries are essential could be made by performing the same experiments with single crystal wires, or with wires having a very large L/R ratio. In such an experiment it would be desirable to check the orientations of different parts of the "single crystals" by x-ray methods, to detect the possible presence of a polycrystalline structure with grain orientations differing by only a few degrees.[14]

Several items may be mentioned which would facilitate the interpretation of future experiments with the usual polycrystalline wires. Grain lengths should be measured over the whole length of the wire, and if possible the distribution of orientation differences between adjacent grains should be checked. Checks should also be made to find out whether the sizes and orientations of the grains change appreciably during the experiment, and whether changes in the creep rate with time are permanent or recoverable. Offsetting and kinking should of course be watched for, though according to the discussion above they are not expected to play a major role in the stress range of most interest.

Diffusional viscosity should not be very greatly affected by small amounts of impurities, once the effect of these impurities on the grain structure has been allowed for. This could easily be checked.

An interesting consequence of the theory of this paper is that the cylindrical surfaces of the wires used in experiments such as those of reference 2 ought to develop shallow depressions along the grain boundaries when they lengthen with time, and ridges when they shrink. This is because the diffusion currents tend to be especially strong at the parts of the surface near the grain boundaries. Use of polished wires would facilitate detection of this effect.

In conclusion, it should be pointed out that the phenomenon of diffusional viscosity, if it turns out to be the predominant cause of creep under very low stresses at sintering temperatures, should be very important in all kinds of sintering processes. This is especially true with regard to the final stages of sintering: the shrinking down of the cavities requires an overall contraction of the specimen, and this can easily be achieved if the crystal grains are free to slide on each other tangentially and to grow or recede normally to their boundaries. In the absence of such processes and of plastic flow within the crystals, the only mechanism which could shrink the cavities would be a diffusion of material from the outer boundary of the specimen, as postulated in the theory of Pines.[15] The latter process obviously gives a rate of densification much slower than the former process, and one which decreases rapidly with increasing dimensions of the specimen.

APPENDIX

In a cubic crystal containing various kinds of atoms, the general expression for the diffusive flux of the ith kind of atom is

$$\mathbf{j}_i=-(n_L/kT)\sum_k D_{ik}\nabla(\mu_k-\mu_h), \qquad (A)$$

where n_L is the number of unit cells in unit volume. Immediately after stresses are applied, the currents of the different species will in general be widely different; however, before any macroscopically recognizable transport of material has occurred concentration gradients will be set up in such directions as to tend to equalize these differences in currents, and if the stresses are small a quasi-steady state will very soon be reached in which $\mathbf{j}_i=X_i\mathbf{j}$, where X_i is the fractional concentration of the ith species and $\mathbf{j}=\sum\mathbf{j}_i$. It will be shown that $\mathbf{j}$ can be expressed in terms of the gradient of a single harmonic function. Equating (A) to $X_i\mathbf{j}$ and solving for $\nabla(\mu_k-\mu_h)$,

$$\nabla(\mu_k-\mu_h)=-(kT/n_L)\mathbf{j}\sum_i (D^{-1})_{ki}X_i, \qquad (B)$$

[14] Such structures have been observed in a "single crystal" of aluminum by P. Lacombe, "Report of a Conference on Strength of Solids" (The Physical Society, London, 1948), p. 91.

[15] B. Y. Pines, J. Tech. Phys., U.S.S.R. **16**, 737 (1946). I am indebted to Dr. G. C. Kuczynski for a translation of this paper.

where D^{-1} is the reciprocal of the matrix D. Now define μ_{0k} at each point of the crystal as the value of μ_k corresponding to the composition occurring at that point but at zero stress. Then by the Gibbs-Duhem relation, $\sum X_k \nabla \mu_{0k} = 0$. Moreover, if the stresses are small ($\mu_k - \mu_k - \mu_{0k}$) will be a small quantity of the first order in the stresses. Multiplying (B) by X_k and summing gives

$$-(kT/n_L)\mathbf{j} \sum_{i,k} (D^{-1})_{ki} X_i X_k = \sum_k X_k \nabla(\mu_k - \mu_k - \mu_{0k})$$

$$\approx \nabla \sum_k X_k(\mu_k - \mu_k - \mu_{0k}) \quad \text{(C)}$$

if we neglect quantities of the second order in the stresses, since ∇X_k is a small quantity of the first order. Thus we can write, in formal analogy to (1),

$$\mathbf{j} = -(n_L D'/kT)\nabla f \quad \text{(D)}$$

with

$$f = \sum_k X_k(\mu_k - \mu_k - \mu_{0k}) \quad \text{(E)}$$

and

$$D' = [\sum_{i,k} (D^{-1})_{ik} X_i X_k]^{-1}. \quad \text{(F)}$$

An argument of the same type as that leading to (2) shows that at the boundary of the crystal

$$f = -p_m \sum_k X_k(\partial v/\partial N_k) = -p_m \Omega_0, \quad \text{(G)}$$

where $\partial v/\partial N_k$ is the partial atomic volume of the kth component and Ω_0 is the mean atomic volume.

The form of the matrix D_{ik} of course depends on the mechanism of diffusion, a subject beyond the scope of this paper. We may merely mention that for purely interstitial diffusion the matrix is diagonal, so that for a two-component crystal

$$D' = (X_1^2/D_{11} + X_2^2/D_{22})^{-1},$$

while for a two-component crystal in which there is no diffusion except by place interchange the determinant $|D_{ik}|$ vanishes, making $D' = 0$, as is to be expected for this model.

Streak Photography

IRWIN VIGNESS AND R. C. NOWAK
Mechanics Division, Shock and Vibration Branch, Naval Research Laboratory, Washington, D. C.
(Received December 5, 1949)

A description is given of some methods of "streak photography" used for obtaining displacement records as functions of time. Ordinary streak photography consists of photographing a moving object on a moving film camera with an open shutter. The motion of the object should be in a plane parallel to the film and perpendicular to its motion. Large errors in velocity determinations may occur if the motion of the object has components parallel to the film motion. A modified method of streak photography is described in which these errors are reduced to small values. By this method time-displacement curves have been recorded for displacements of many inches and periods of several seconds with accuracies within a few thousandths of an inch and a few hundred thousandths of a second.

INTRODUCTION

THE measurement of transient motions of mechanical elements is of constant necessity in laboratory and field problems associated with mechanical shock and vibration. The conditions of the experiments require methods of measurements that are easy to perform and that do not appreciably load the mechanical parts under study. While high speed moving pictures provide this together with a wealth of other information, the work of transcribing the motion picture records into quantitative displacement curves is too laborious for many conditions. Methods of streak photography are easy to perform and can often be done with ordinary laboratory equipment. Many types of oscillographic recording cameras can be used with little or no modification. Several Naval Laboratories have made extensive use of the methods of streak photography to conveniently obtain accurate displacement curves.[1] However, no descriptions have been made of these methods except as incidental to other experiments. This paper describes the methods of streak photography most generally used.

[1] Vigness, Kammer, and Holt, "Shock and vibration instrumentation and measurements," Second Partial Report, Naval Research Laboratory Report No. 0-2645 (September, 1945).

ORDINARY STREAK PHOTOGRAPHY

One of the simplest methods of recording transient motions of an object is to photograph the object with a moving film camera. The film of the camera is made to move at a known speed and the shutter is held open during the entire exposure. The camera is arranged so that the direction of motion of the object under study is perpendicular to the camera optic axis and to the direction of film motion. Sufficiently good contrast can often be obtained without any special preparation of the object, but usually it is advisable to arrange bright spots with a dark background on the object whose motion is being studied. These bright spots may be dots of white paint, ball bearings, or small luminous sources. Ball bearings are particularly advantageous if very fine lines are desired. A small light source some distance from the ball should be used, the reflection of this light by the ball closely approximates a point. The ball should be small so that the reflected light will continue to come from essentially the same location on the ball. Many spots may be in the field of view, and, consistent with the requirements of depth of focus, they can be spread over any distance from the camera. Of course the displacement scale will depend upon this distance. Illumination

FIG. 1. Ordinary streak photograph showing the horizontal components of a displacement of a hammer and table when the table is struck by the hammer.

can be conveniently obtained for the non-luminous spots by photo-flood lights of the R3 type.

An example shown in Fig. 1 illustrates a hammer striking a steel table. The hammer swung as a pendulum and struck the edge of the steel table which was mounted on rollers and was constrained to move in but one direction. The upper part of the figure shows the motion of the hammer. The impact occurs at the sudden change of slope. The velocity of the hammer and table before and after impact can be determined from the slopes of the corresponding streaks. The boundary streaks of the hammer and table do not come into contact at impact as the camera did not observe the immediate areas of impact.

Special methods of streak photography have been applied for measuring the rate of flame propagation and for determining the velocity of the luminous areas of explosions. A precise and very accurate method of determining the deceleration of a projectile, and hence the force exerted by the projectile, is routinely performed by the Ballistics Branch of the Naval Research Laboratory by an unusual streak photographic technique.[2] In this case the moving film is contained on the inside of a drum that is 4 inches in diameter and that rotates so that film speeds up to 600 feet per second are attained. An intense source of light is flashed on for a half-millisecond during the passage of the projectile and the shadow of one point of the base of the projectile is recorded on the moving film.

The method of streak photography results in a direct plot, in Cartesian coordinates, of the displacement time history of as many locations as can be contained in the field of view. It possesses the advantage over high speed moving pictures of greater accuracy and the elimination of point by point plotting of curves when displacement records are required. High speed moving pictures, of course, possess many advantages in other respects.

Accurate time scales can be obtained by a variety of methods. Perhaps the most obvious is to place a stationary arc in the camera field of view. "Point" arc light sources similar to that supplied by Central Scientific Company can be controlled by an amplified electrical signal that has been generated by a tuning fork. When a camera slit is used, as for the modified method of streak photography described below, it is convenient to reflect the light from a "Strobolux" (General Radio Company) into the slit so that timing lines are formed at regular intervals. Some oscillographic cameras use spark discharges that expose the film adjacent to the spark. Others have an intermittent light source internal to the camera and focused on the film. Rotating drum cameras may have their film speeds determined simply by measuring the rotational speed of the drum. Generally, little difficulty is encountered in obtaining a precise time scale even if the film speed is non-uniform.

MODIFIED STREAK PHOTOGRAPHY

A source of error in the ordinary method of streak photography lies in the possibility of small motions of the object in a direction parallel to the film motion. In order to eliminate this error, to measure large displacements (4 or 5 inches), and to obtain an accuracy within a few thousandths of an inch, a modified method of streak photography was devised. The principles of this method are shown in Fig. 2. In this figure the object consists of a series of bright lines on a dark background. The lines are contained in a plane parallel to the direction of the film motion. The direction of motion of the object, which is rigidly attached to the part whose motion is being studied, should be perpendicular to the camera optic axis and to the direction of film motion. The camera lens will focus an image of these lines on the film. A slit and a cylindrical lens (the lens is optional) placed near the film focuses the line to an approximate point. Motions of the object in a direction parallel to the film motion will not result in any displacement of the image on the film. Small rotations in a plane perpendicular to the camera axis will produce second-order errors, and motions out of this plane will result in the usual scale changes and depth of focus problems. Such an arrangement, used at the Naval Research Laboratory, had an object that consisted of a grid of music wire; the wires of which were about 0.01 of an inch in diameter and were spaced 0.200

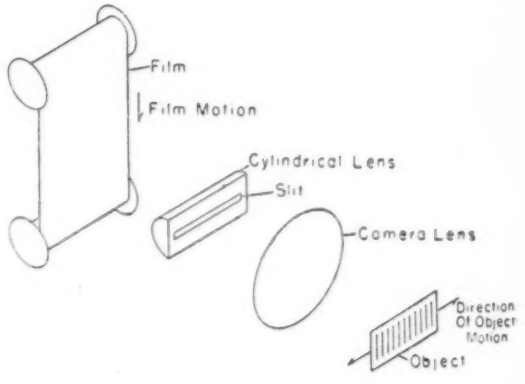
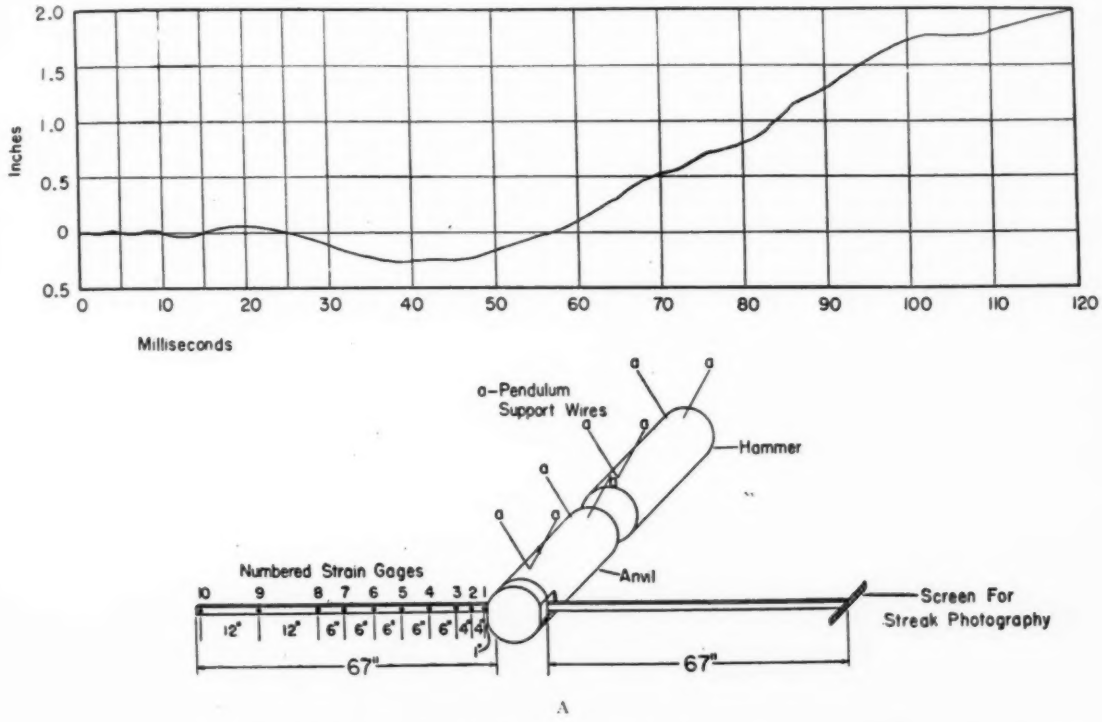


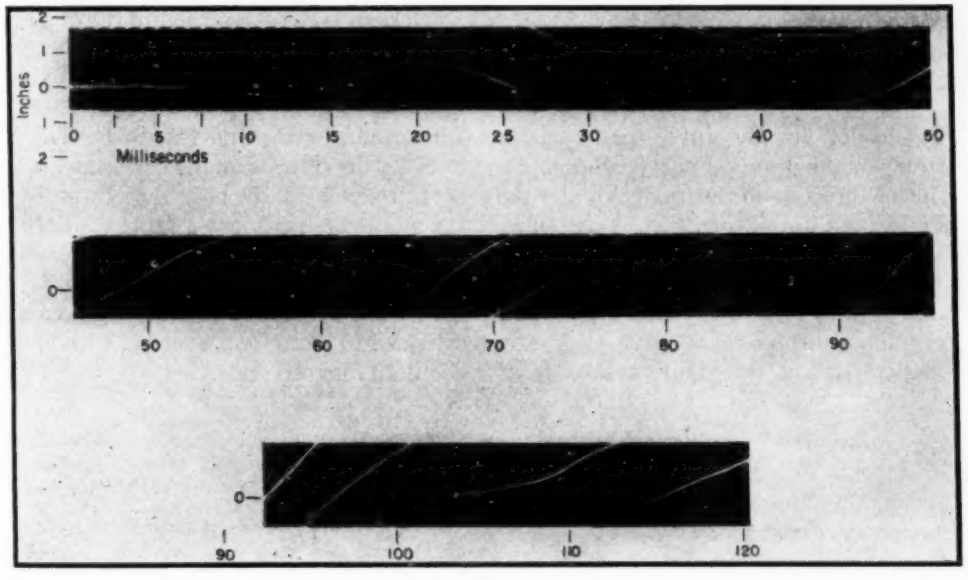
FIG. 2. Principal parts used for the modified method of streak photography. The slit should be near to the film compared with the film-lens distance.

[2] A. Victor Masket, "The measurement of forces resisting armor penetration," J. App. Phys. **20**, 132–140 (February, 1949).

of an inch apart in narrow grooves that were machined in a black Bakelite base. The object was varnished so as to cement the wire in position. The use of a wire for the modified method of streak photography is analogous to the use of a ball for the ordinary method and makes possible a very fine line image. Reasonably good results were also obtained by simply filling sharp machined scratches with white ink in a similar Bakelite base. A convenient camera arrangement consists of a 2-inch focal length lens attached to a moving film oscillograph camera (General Radio GR 651A-E). An extension tube was usually used that held the lens about 9 inches from the film. This arrangement caused the object to be enlarged about 4 times in linear dimensions and the image of 2 object wires could be seen on the 35-mm film when one image was near the edge of the film. This fixed the displacement scale on the film. A cylindrical lens was cemented over a $\frac{1}{16}$-inch wide slit which replaced the



A



B

FIG. 3. Displacement curves obtained by streak photography. The upper curve was derived from the photographic record shown on part B of the figure.

ordinary film gate. Fair pictures may be had with a $\frac{1}{32}$-inch slit and no cylindrical lens. When wires were used for the grid lines the best results were obtained with a small source of illumination, such as a laboratory arc with a focusing lens. The light must of course be in a plane perpendicular to the grid lines and containing the lens optic axis. Sufficient illumination is easily obtained for writing speeds up to 11 ft./sec.

Figure 3 illustrates this method as used to determine the motion of the free end of a cantilever beam when the built-in end suddenly acquires and maintains a uniform velocity. The experimental arrangement is shown in the center part of this figure. The hammer strikes the anvil

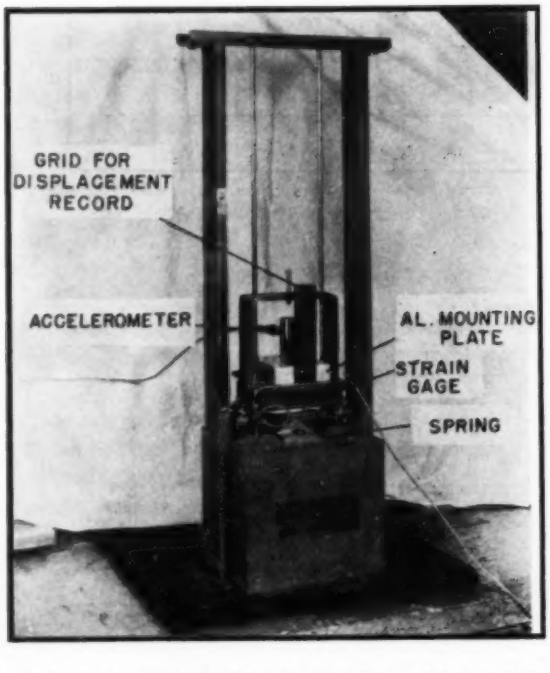


FIG. 4. An instrument shock testing machine with attached grid for streak photography.

and imparts a velocity to the anvil that remains essentially constant for the duration of the experiment. The velocity is in the direction of the anvil axis and the displacement is measured in this direction. A reproduction of the streak photograph is shown on the lower part of Fig. 3. The displacement scale is determined by the vertical separation of lines which represents 0.2 inch on the object. A continuous displacement curve, such as is shown on the upper part of this figure results from

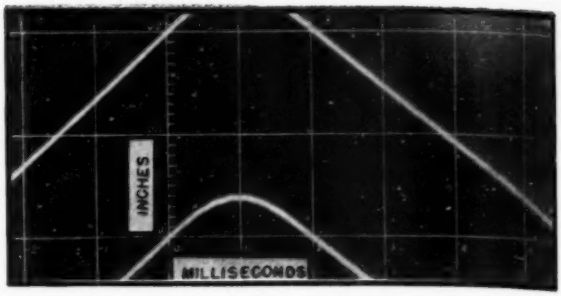


FIG. 5. Streak photograph illustrating the motion of the platform of the instrument shock testing machine.

joining together the various curves, shown on the lower section, with proper multiples of the 0.2-inch vertical separation.

Figures 4 and 5 show the application of the modified method of streak photography for the determination of the motions of a shock testing mechanism. As shown on Fig. 4 a small platform falls on a stiff spring and rebounds. The vertical motion of this platform is shown through the impact time in Fig. 5. The object grid in this case consisted of white inked grooves that were machined at 0.2-inch intervals.

SUMMARY AND CONCLUSIONS

Ordinary streak photography which consists of photographing a moving object on a moving film shutterless camera is a convenient and accurate method of obtaining displacement curves as a function of time provided the direction of motion studied is previously known and is unidirectional. When there are components of motion, which result in image displacements parallel to the motion of the recording film, a modified method of streak photography must be used. In this method a small rigid body is attached to the object whose motion is being studied. One or more parallel lines, which are made perpendicular to the principal component of motion and which are contained in the focal plane of the camera, are scribed on this body. The image of these lines on the camera film is cut to an approximate point by introducing a slit near the film or by introducing a slit and cylindrical lens a short distance from the film. By the above method transient displacements of many inches can be measured to an accuracy within a few thousandths of an inch. The time scale can easily be extended to more than a second with an accuracy up to about 10 microseconds.

Surface Tension Effects in Oriented Thin Silver Films*

G. W. JOHNSON†
Department of Physics, State College of Washington, Pullman, Washington
(Received November 7, 1949)

Structural changes of thin oriented films of silver produced by condensation in a vacuum on cleaved surfaces of single crystals of rocksalt were followed as a function of temperature. Marked agglomeration occurred in films of thicknesses less than 1000A. For thicker films in the 1000–2000A range rectangular holes appeared in the films with their edges aligned with the 110 directions of the rocksalt. The effects are attributed to the readjustment of the oriented surface under the influence of surface tension forces.

THE behavior of thin films of silver condensed in a vacuum on various substrates has been the subject of numerous investigations. A large number of these investigations have been concerned with the measurement of the electrical resistivity and its temperature dependence. The general features of all of the resistance measurements have been discussed by Appleyard.[1] There are two principal results common to all of the resistance measurements. The first is that the resistivity of thin films is greater by factors of 10^3 to 10^7 than that of the bulk metal. The second is the irreversible increase of resistance with increase of temperature. Moreover, for films of a given thickness the resistivity is greater the higher the temperature of the substrate at the time of deposition. These observations led to the view that the film as deposited has a granular structure which is a result of the agglomeration of the metal under the surface tension forces. That such agglomeration occurs has been shown directly by the work of Estermann,[2] and by that of Andrade and Martindale.[3] Recent photographs made with the electron microscope clearly reveal the grained structure of thin metallic films.[4] The energetics and proposed mechanisms of aggregation are given by Appleyard and by Lennard-Jones.[5] It was suggested that aggregation occurs if the attractive forces between the condensed metallic atoms is greater than the forces between the metallic atoms and those of the substrate. The tendency to aggregation is opposed by the periodic field of the surface and aggregation can only occur if the rate of surface diffusion at the given temperature is great enough. It is apparent that an increase of temperature will favor the tendency to aggregation since higher temperature results in more rapid diffusion rates.

In the present work the structural changes of oriented thin films of silver all condensed on rocksalt at 200°C were followed as they were heat-treated subsequently at higher temperature. The films were produced by vacuum distillation of silver on freshly cleaved crystals of rocksalt preheated to about 200 C in accordance with well-known techniques.[6] Wilman and Goche[7] in a detailed study of the structure of such films showed them to possess a twinned structure as grown but heat-treatment at 500 C for a few minutes converted them to relatively perfect single crystals.

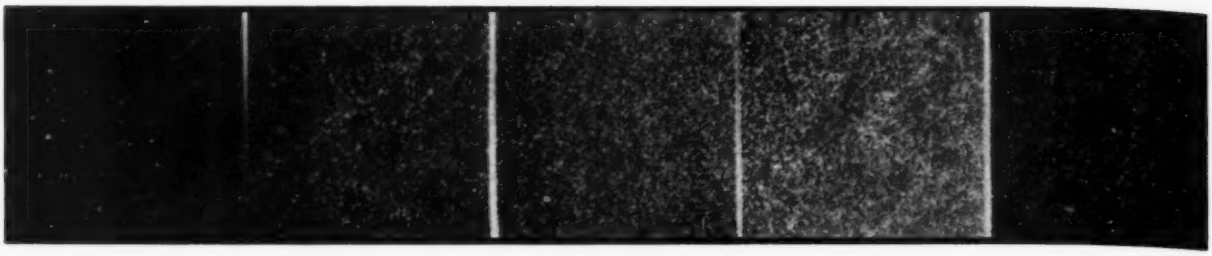
EXPERIMENTAL METHOD

The method of Lassen and Bruck was used to produce thin oriented films of silver. A single crystal of rocksalt was cleaved and was introduced into the tube which was immediately evacuated. At a pressure of less than 0.01 micron and after preheating the rocksalt crystals to about 200 C, spectroscopically pure silver was vaporized from a tungsten filament and condensed on the rocksalt. The vacuum chamber was designed to permit the introduction of five cleaved crystals with dimensions of about 4×4 mm. The crystals were supported on a glass slide movable at will beneath the vaporizer by means of an external magnet. This arrangement made it possible to vary the exposure to the beam of silver. By using different exposure times the film thickness could be varied for the same heating current and for the same geometry.

The films thus produced were examined under vertical illumination with both bright field and polarized light (crossed nicols) using a Bausch and Lomb metallograph. For the films used in this work it was found in general that polarized lighting was more effective in revealing the structure than was bright field illumination. Accordingly most of the photo-micrographs herein reproduced were made under polarized illumination. The electron diffraction attachment of an RCA electron microscope was used to determine the orientation of the films.

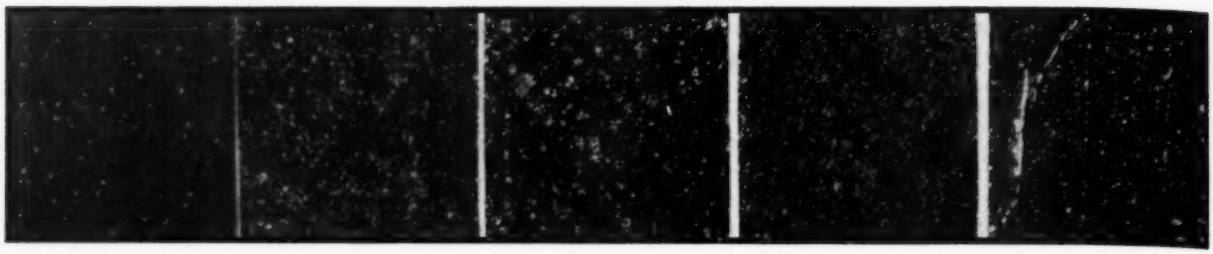
After inspection, the films still supported on the rocksalt substrate were introduced into a vacuum furnace for heat-treatment. The films were heated to various temperatures in the range of 300–500 C for times of from two to twelve hours. After each such

* This work was done under contract with ONR.
† Now at Brookhaven National Laboratory, Upton, New York.
[1] E. T. S. Appleyard, Proc. Phys. Soc. **49**, Extra Part, 118 (1937).
[2] I. Estermann, Zeits. f. physik. Chemie **106**, 403 (1923).
[3] E. N. da C. Andrade and J. G. Martindale, Phil. Trans. Roy. Soc. **235**, 69 (1935).
[4] R. G. Picard and O. S. Duffendack, J. App. Phys. **14**, 291 (1943); H. Levinstein, J. App. Phys. **20**, 306 (1949).
[5] J. E. Lennard-Jones, Proc. Phys. Soc. **49**, Extra Part, 140 (1937).
[6] H. Lassen, Physik. Zeits. **35**, 172 (1934); F. Kirchner and H. Lassen, Ann. d. Physik **24**, 113 (1935); L. Bruck, Ann. d. Physik **26**, 233 (1936); W. Cochrane, Proc. Phys. Soc. **48**, 723 (1936); V. O. Rudiger, Ann. d. Physik **30**, 505 (1937).
[7] O. Goche and H. Wilman, Proc. Phys. Soc. **51**, 625 (1939).



(a) (b) (c) (d) (e)

FIG. 1. Silver film. Thickness~400A. Polarized light crossed nicols 500×. a. As deposited, b. 320°C; 4 hours, c. 400°C; 4 hours, d. 475°C; 4 hours, e. 500°C; 4 hours.



(a) (b) (c) (d) (e)

FIG. 2. Silver film. Thickness~700A. Polarized light crossed nicols 500×. a. As deposited, b. 320°C; 4 hours, c. 400°C; 4 hours, d. 475°C; 4 hours, e. same as (d) but bright field illumination.

treatment the films were examined optically to note any changes in structure. The nominal film thicknesses were calculated from a chemical determination of the silver, using a colorimetric method, the apparent area of the film, and the bulk density. The nominal thickness is the thickness a film would have if it were uniformly distributed in depth over the surface.

RESULTS

Films were prepared in thicknesses extending from 100A to about 4000A. All were heat-treated in a vacuum (less than 0.01 micron) at temperatures between 200 C, the temperature at which they were grown, and 530 C. Above 500 C the rocksalt vaporized profusely with the result that the films were ultimately detached from the substrate and destroyed. It was found that the behavior of the films depended strongly on their thickness as might have been anticipated. For this reason it is convenient to discuss the results in terms of four different ranges of thickness as indicated in the following. All of the films to be described gave diffraction patterns characteristic of single crystals. All photographs reproduced here were made with the cube edges (100 directions) of the rocksalt parallel to the edges of the photographs.

100–500A.—Typical structures of films of this thickness are shown in Fig. 1. These films are noted to have a granular appearance. At higher magnification they are found to possess a network structure which is characteristic of these films and has been observed and reported by many investigators. Recent papers by Levinstein[4] with indium and antimony, and by Picard and Duffendack[4] with aluminum, gold, and copper for unoriented metallic films deposited on collodion, include photo-micrographs of the network structure obtained with an electron microscope. In their work the aggregation of metal into local regions separated by thinner regions or canals is clearly shown. The thinnest films in this thickness range were non-conducting as was indicated when attempts were made to get electron diffraction patterns. Under such conditions it is impossible to get a diffraction pattern because the beam drifts about capriciously as local regions acquire or lose electric charge. The granular structure may be seen on the film as deposited (Fig. 1a). In this case the individual grains are of the order of 1 micron in diameter and on the photo-micrograph (500×) are about 1 mm in diameter. However, the structure becomes accentuated with heat-treatment for a few hours at 320 C (Fig. 1b). There was little observable change in the films on heat-treatment at higher temperatures until a temperature of 500 C was reached, at which point the aggregation of the film became marked (Fig. 1e).

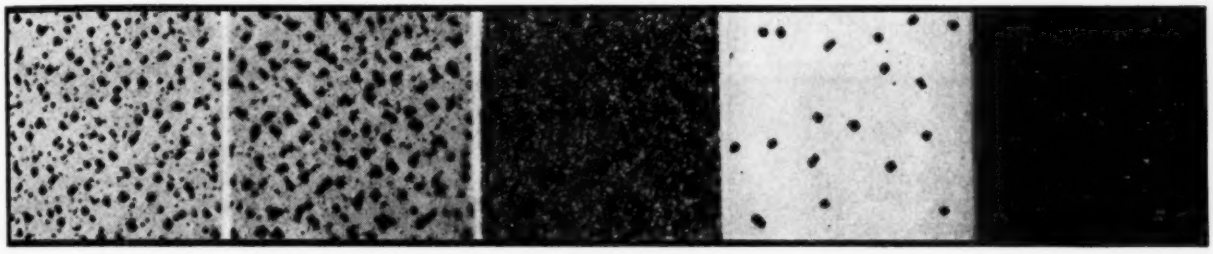
Occasionally the silver as deposited showed a preference for cleavage lines on the crystal. As the temperature was raised the lines faded out. These effects are illustrated in Figs. 1a and 1b which are photographs of the same area before and after heat-treatment.

500–1000A.—The structures observed in this thickness range and their change with temperature can be seen in Fig. 2. The films as deposited resembled those of the thinner range but their changes with temperature are different. On heating to 320 C for four hours the network structure became plainer, with some indication of agglomeration. After heat-treatment at 400 C the

network structure practically disappeared and the silver had collected into relatively large aggregations. At 475 C in some of the films the silver collected along cleavage lines to form ridges and also in small stubby rods extending normal to the surface. The protuberances of silver, whether along cleavage lines or in isolated regions, seemed to grow to about the same height above the crystal surface. The height of growth was about 0.001 mm and was measured by noting the movement of the microscope objective to bring into focus alternately the tops of the growths and the surface of the film.

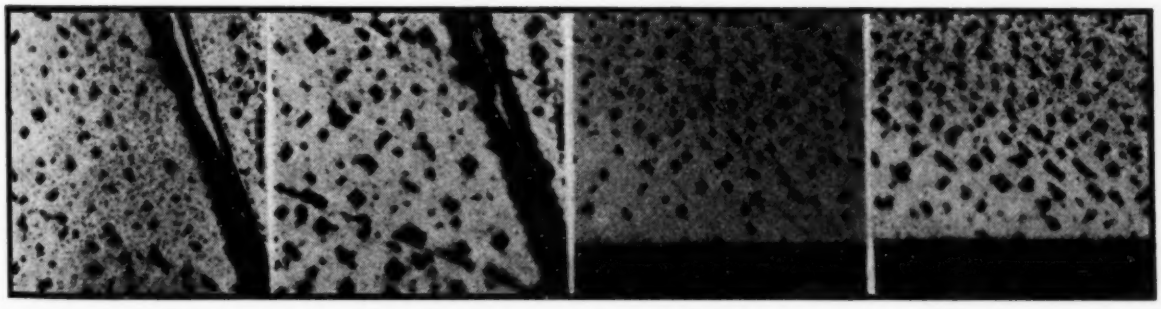
1000–1500A.—While films of thickness less than 1000A all underwent some adjustment on heating to 320 C, films of greater thickness did not change in appearance until the temperature was increased at least to 400 C. In this thickness range the thinner films on heating to 420 C for four hours were perforated by approximately rectangular windows with the edges aligned with the (110) directions of the rocksalt. The holes were uniformly distributed and were all about the same size within perhaps a factor of two. On increasing the temperature to 450 C the holes increased in size by some 20 percent which was accompanied by the addition of spurs in some cases resulting in changes in shape. In Figs. 3a and 3b are shown some typical holes and also their growth with temperature. Both 3a and 3b were made under bright field illumination and are of the same area of the film. Figure 3c is the same area under polarized light. In this latter photograph the advantage of polarized light for revealing the granular structure is clearly seen, for while the holes are outlined under both bright field and polarized light, the granular structure is only shown under polarized light. For thicker films in this region the rectangular holes appeared also but their density was not as great and the film in general retained its continuity (Figs. 3d and 3e). Further examples of the holes and their growth are shown in Figs. 4a and 4b. Of interest in these photographs is the growth of serrations or steps from a scratch in the film. In Figs. 4c and 4d are shown holes in the neighborhood of a cleaved edge of rocksalt to show the relative orientation of the holes with respect to the (100) directions of the rocksalt. The preference of the straight edges of the windows for the face diagonal directions is clearly illustrated. The density of the windows decreases as a scratch or a free edge is approached. It is of interest also to note that the windows or parts thereof seem to develop from such free edges. (See scratch in Fig. 4b.)

1500–2000A.—Films of this thickness were quite stable and did not show any marked structural changes until heated to 480 C, whereupon the films were found to be criss-crossed by systems of straight lines spaced about 1 micron apart and directed along the (110) directions of the rocksalt. The lines became more pronounced on further heating at 520 C until finally the films disintegrated. Windows sometimes appeared in the form of crosses or rectangular openings with serrated edges. The serrations appeared to have magnitudes corresponding



(a) (b) (c) (d) (e)

FIG. 3. Silver film. a. Thickness~1100A. 4 hours at 420 C. Bright field 500×, b. same area as (a) after 4 hours at 450 C. Bright field 500×, c. same condition as (b) but polarized light, d. thickness 1400A. 4 hours at 480 C. Bright field 500×, e. same as (d) but polarized light.



(a) (b) (c) (d)

FIG. 4. Silver film. Bright field~500×. a. Appearance of serrations near scratch on film. 420 C, b. growth of holes and serrations at elevated temperature 480 C, c. orientation of holes with respect to cleaved edge of rocksalt substrate 420 C, d. growth of holes at 480 C.

to the observed spacing of the lines covering the entire film. These effects are all shown in Fig. 5. On the original photo-micrographs could be seen systems of faint straight lines running along the face diagonals of the rocksalt.

Above 2000A.—Films thicker than 2000A did not undergo any pronounced structural changes with heat-treatment until they disintegrated after heating to 520 C for eight hours. The heaviest film studied was about 4000A in thickness.

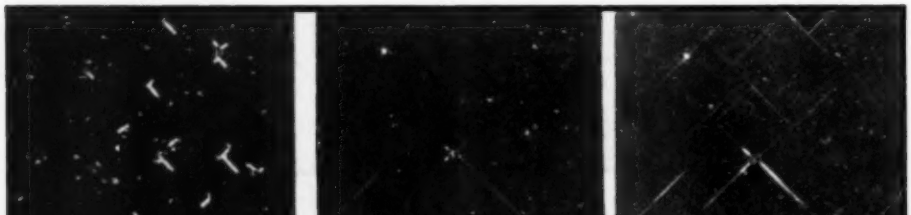
GENERAL OBSERVATIONS

All of the films as deposited gave bright specular reflections. The mirror surfaces were maintained until marked agglomeration occurred or until holes formed in the films at which point the films assumed a matte appearance.

On high rates of deposition the silver would occasionally collect in globules along lines of conchoidal fracture on the cleaved surface of the rocksalt. This effect is illustrated in Fig. 6a.

By exposing the films to a moist atmosphere small sodium chloride crystals could be produced on top of the silver. Such crystals were found to grow over a period of a few months while the films were being stored in a calcium chloride desiccator. They can be readily produced by breathing lightly on a film supported on rocksalt. In Figs. 6b and 6c are shown some of these small crystals. In this case the salt crystals are all aligned with edges parallel to the 110 direction of the substrate and also along the 110 direction of the silver, although often they are rotated 45° with respect to this direction. In the lower right-hand corner of Fig. 6b may be noted one crystallite rotated 45° with respect to the others. These two orientations were the only ones observed on any of the films. Since the silver as deposited was oriented with its structure parallel to that of the rocksalt, then the same relative orientation would be expected if sodium chloride were deposited on silver. The salt on silver should then have the same orientation as the underlying substrate. In the photographs the 100 directions of the rocksalt substrate are parallel to the edges of the photograph. If we assume that the cube edges of the sodium chloride crystallites are the 100 directions then the crystallites are rotated 45° with respect to the expected orientation. This orientation is consistent with that reported by Brück and confirmed by Goche and Wilman for films which were not thoroughly rinsed diffraction points were observed and were attributed to sodium chloride retained on the film with the same orientation as was found here.

To determine whether the epitaxy noted in the foregoing for sodium chloride on an oriented silver film was the result of an influence exerted by the substrate through small pores, a film was floated off the rocksalt in water and picked up on a degreased silver plate.[8] After drying, the film was rinsed in a saturated aqueous solution of sodium chloride and quickly dried. Small crystals of sodium chloride were thus produced on the silver film. Examination under the microscope showed the crystallites to be aligned and with the same orientation as was observed for the films on the rocksalt substrate (Fig. 7). This result is of interest in that it gives a ready means of getting a rough check on the degree of orientation of a



(a) (b) (c)

FIG. 5. Silver film. Polarized light (crossed nicols) 500×. a. Thickness 1600A 480 C for 4 hours. b. Thickness 1800A 480 C for 4 hours. c. Same area as (b) after 4 hours at 520 C.

(a) (b) (c)

FIG. 6. a. Aggregation of silver on cleavage lines. Polarized light 500×. b. Rocksalt crystals on silver showing epitaxy. Bright field 500×. c. Same as (b). Polarized light 500×.

[8] The author is indebted to Mr. Theodore Watanabe, Washington State College, who worked out the technique of mounting the films.

film without resorting to electron diffraction. Indeed, it may be more sensitive to the orientation in the first layer since the surface forces are probably of short range.

DISCUSSION

The network or granular structure which was observed for the thinner films is characteristic of such films and is well-understood. The metal as deposited agglomerates into small regions under the influence of surface tension forces which for very thin films may be very great. On increasing the temperature further agglomeration may take place. In the present work this effect is noted in the thinnest films where the network structure is accentuated as the temperature is increased and finally the film breaks into small isolated agglomerated regions.

As the film thickness is increased the adjustment of the surface under the effects of the surface tension forces are not as drastic. For films of intermediate thickness this leads to the formation of holes which are evenly distributed and have sides parallel to the 110 direction of the rocksalt. The origin of the holes is suggested in the following. The silver film has the same orientation as the rocksalt and therefore the lines of closest packing of silver atoms are in directions parallel to the face diagonals of the rocksalt. These directions are therefore directions along which the binding between silver atoms is greatest. As the temperature is raised the film is put under increasing tension as a result of the greater thermal expansion of the rocksalt than that of the silver. At some temperature, depending upon the film thickness, the film fails and since it apparently fails along lines of closest packing the process of failure must be one of shear. The surface tension forces then cause the film to contract further. The difference between the coefficients of expansion can only account for about 1 percent total strain. Since the holes are much larger than this, one concludes that the primary cause of growth of the holes lies with the surface tension forces. The holes continue to grow somewhat with increase of temperature. For the thickest films (1500–2000A) the surface tension forces are not sufficiently great to cause large changes in the surface. The film is strained enough, however, to lead to rupture or local yielding sufficient to be visible under the microscope. The lines broaden on further heating indicating an extension of the rupture process but also indicating the films are sufficiently thick to resist any large scale readjustment to the surface tension forces.

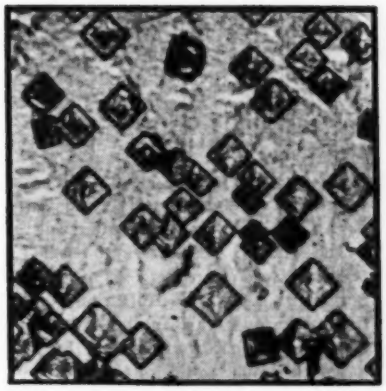


FIG. 7. Sodium Chloride Crystals deposited from an aqueous solution on an oriented silver film supported on a silver plate 400×.

Essig[9] noted in one case the formation of regular holes in a silver film condensed on mica. The windows were bounded by lines intersecting at 60° and 120°. The explanation of the origin of these holes must be similar to that given here for rectangular holes. The silver for the conditions used by Essig must have condensed with the 111 plane parallel to the surface. For this orientation there are three lines of densest packing of atoms in the plane of the surface and they intersect at 120°. Thus when failure occurs if it were by a shearing process it would be expected to take place along lines of closest packing, in which case the windows would have straight sides intersecting at the observed angles. Since the relative orientation of the silver relative to the mica was not determined, it is not known whether the orientations would be consistent with the picture postulated here. However, there appears to be little question about the general description of the result.

ACKNOWLEDGMENTS

The author is indebted to Mr. Mark Adams, Washington State Institute of Technology, for the colorimetric determinations, to Mr. R. F. Smith of Brookhaven National Laboratory, for photo-micrography (Fig. 7), and to Dr. A. W. McReynolds for many helpful criticisms.

[9] S. F. Essig, J. App. Phys. **10**, 61 (1939); Phys. Rev. **55**, 229 (1939).

Upper and Lower Bounds of Eigenvalues for Composite-Type Regions

E. L. Chu
Microwave Laboratory, Stanford University, California
(Received October 21, 1949)

A survey is made to compare the essential features of different variational methods. Schwinger's method is found to have special merits for boundary value and eigenvalue problems concerning composite-type regions. It can give both an upper bound and a lower bound for the eigenvalue of any mode, one bound being as good as the other. It closely resembles the Courant-Trefftz method and is simpler. In both methods the trial functions are so chosen as to satisfy the differential equation while the boundary or continuity conditions are perturbed. The applicability and power of Schwinger's method, being hitherto demonstrated only for particular examples, are exhibited for a whole class of problems by a general formulation of the method in precise terms. The formulation for the upper-bound case presents no difficulty, and a simple proof is given in this paper. A similar proof for the lower-bound case is possible but has a more restricted field of application. The general formulation places the two different cases on equal footing. A way of avoiding the difficulty of having no exact knowledge of the lower eigenfunctions in higher mode cases is discussed in detail.

I. SURVEY OF METHODS

THE best-known variational method of Rayleigh[1]-Ritz[2] has the great drawback of giving only an upper bound for the true eigenvalue with no information about the accuracy of the bound. One is usually content to indicate the degree of accuracy obtainable by applying the method to problems which can be solved exactly and comparing the approximate value with the true one. This is not quite satisfactory, because the examples chosen may happen to be particularly favorable. Besides, the minimizing process of Ritz is somewhat involved especially for high orders of approximation or for higher modes. There have been recommended, notably by Galerkin,[3] some modifications of the Ritz method, in which the minimizing process is simplified. Yet, the character of overestimation remains the same.

Weinstein[4] has devised an ingenious method which is capable of giving both an upper and a lower bound for the eigenvalue for any one mode. It is rather unfortunate that the expression to be minimized is somewhat complicated, being the difference between a certain integral and the square of another, and the evaluation of one of these integrals is considerably more difficult than that of the Rayleigh-Ritz expression. To facilitate the evaluation of the integrals, only simple trial functions can be used, thus limiting the accuracy of the process. Though interesting and attractive, this method has not been used much in practice.

The suggestion that a lower bound for the eigenvalue can be obtained by calculating the stationary value of the variational expression under loosened boundary conditions or loosened conditions of continuity was first made by Courant.[5] The value calculated must be stationary with respect to small arbitrary variations for otherwise it may not be a lower bound. How this can be done was first shown by Trefftz.[6] Unlike the Rayleigh-Ritz method, the trial functions used by Trefftz satisfy the differential equation but are subject to less restrictive boundary and continuity conditions. The Courant-Trefftz method may be considered as the counterpart of the Rayleigh-Ritz method, because the requirements on the trial functions for these two methods are just opposite to each other. In view of the elegance and the wide applicability of the Courant-Trefftz method, it seems desirable to give here a brief account of its application.

Consider the Sturm-Liouville differential equation for two independent variables x and y:

$$L[u]+\lambda\rho u=(pu_x)_x+(pu_y)_y-qu+\lambda\rho u=0, \quad (p>0,\ \rho>0), \quad (1)$$

for a region R with boundary Γ.[7] The boundary condition is given as $u=f(s)$. According to Trefftz, the problem is to be modified by replacing the given boundary condition $u=f(s)$ with a set of weaker ones:

$$\int_\Gamma \{w-f(s)\}g_i(s)ds=0, \quad i=1, 2, \cdots, m,$$

where $g_i(s)$ are arbitrary functions chosen in such a way that the differential equation $L[w]+\mu\rho w=0$ can be solved exactly. As in the Rayleigh-Ritz method, the function w that corresponds to the lowest eigenvalue μ will minimize the expression

$$D[w]=\int_R\int p(w_x^2+w_y^2)dxdy+\int_R\int qw^2dxdy, \quad (2)$$

but under a set of modified boundary conditions. The

[1] Lord Rayleigh, Phil. Trans. **161**, 77 (1870); *Scientific Papers*, (Cambridge University Press, London, 1899), Vol. 1, p. 57.
[2] W. Ritz, *Oeuvres* (Gauthier-Villars, Paris, 1911), pp. 192–316.
[3] See F. Pfeiffer, Handbuch der Physik **6**, 345 (1928); also Biezeno-Grammel, Tech. Dynamik (Verlag. Julius Springer, Berlin, 1939), 167–169 (1939).
[4] D. H. Weinstein, Proc. Nat. Acad. Sci. **20**, 529 (1934); see also J. K. L. MacDonald, Phys. Rev. **46**, 828 (1934).
[5] R. Courant, Math. Ann. **97**, 711 (1927).
[6] E. Trefftz, Math. Ann. **100**, 503–521 (1928).
[7] In Eq. (1), p, q, and ρ are functions of x and y while λ is a constant. $p>0$ and $\rho>0$ in R. $L[u]=(pu_x)_x+(pu_y)_y-qu$ is simply a definition.

normalization relation

$$\int_R\int \rho w^2 dxdy = 1$$

is understood.

Since the modified conditions are weaker than the given condition, that is, the set of compatible functions for w contains the set of functions for u, so the minimum value of $D[w]$ can never be greater than that of $D[u]$. Thus $\mu \leqq \lambda$, μ, and λ are the minimum values of $D[w]$ and $D[u]$, respectively.

The modified problem is the same as that of finding a function w for which the following functional $I[w]$ is stationary:

$$I[w] = D[w] - \mu \int_R\int \rho w^2 dxdy - 2\sum_1^m c_i \int_\Gamma \{w - f(s)\} g_i(s) ds, \quad (3)$$

where c_i's are Lagrangian multipliers. This can be shown by putting the first variation of $I(w)$ equal to zero, to obtain

$$\int_R\int \{L[w] + \mu\rho w\}\, \delta w dxdy - \int_\Gamma \left\{ p\frac{\partial w}{\partial n} - \sum_1^m c_i g_i(s) \right\} \delta w ds = 0.$$

The volume integral vanishes if w satisfies the equation $L[w]+\mu\rho w=0$. The surface integral will also vanish for any arbitrary δw if

$$p\frac{\partial w}{\partial n} = \sum_1^m c_i g_i(s)$$

on Γ. These requirements on w can be fulfilled if w is taken as

$$w = \sum_1^m c_i w_i(x, y)$$

and $g_i(s)$ as

$$g_i(s) = p(\partial w_i(x, y)/\partial n),$$

where the w_i's are any m particular solutions of the differential equation $L[w]+\mu\rho w=0$.

The $m+1$ constants, μ and the c_i's, are to be determined from $m+1$ simultaneous equations, m of which are the above boundary conditions while the other one is the normalizing relation. As shown above, μ is a lower bound for λ.

The lower bound may be improved by imposing more stringent boundary conditions. In fact, as m approaches infinity, w may be made to approach the true eigenfunction. Thus μ will approach λ. In some cases it is desirable not to relax the boundary conditions but the conditions of continuity. Such cases have also been considered by Trefftz.

In passing, we may remark that the Trefftz method can also be used for calculating upper bounds. What we have to do is simply to make the boundary conditions more stringent than given in the original problem. However, there would be no advantage to do so even where the trial functions can be easily chosen, because the algebraic process of the Trefftz method is still more involved than that of Ritz.

One method other than Ritz's for calculating upper bounds needs special mentioning. This was described by Grammel[8] and may be considered as the counterpart of Ritz's method or rather that of Galerkin's method, from the viewpoint that this method is more intimately connected with the integral equation rather than with the differential equation.

To be specific, if the differential and integral equations for the same problem are given as $L[u]+\lambda\rho u=0$ and

$$u(x, y) = \lambda \int_R\int G(x, y; x', y'; 0) \times u(x', y')\rho(x', y')dx'dy' \quad (4)$$

respectively, where $G(x, y; x', y'; 0)$ is the appropriate Green's function of the equation $L[u]=0$, then the Grammel method deals with the equation

$$\int_R\int \left[w(x, y) - \lambda \int_R\int G(x, y; x', y'; 0) w(x', y') \times \rho(x', y')dx'dy' \right] \rho(x, y)\delta w dxdy = 0 \quad (5)$$

while Galerkin's method deals with

$$\int_R\int \{L[w] + \lambda\rho w\}\, \delta w dxdy = 0. \quad (6)$$

In both methods we take

$$w(x, y) = \sum_1^m c_i \phi_i(x, y) \quad (7)$$

as the trial function. The functions ϕ_i are any suitable coordinate functions satisfying the appropriate boundary conditions of the problem. We restrict ourselves only to such variations as

$$\delta w = \sum_1^m \phi_i(x, y)\delta c_i.$$

The variational problem is then reduced to the solution

[8] R. Grammel, Proc. 5th Intern. Congress for App. Mech. 691–693 (1938); also Biezeno-Grammel, reference 3.

of m simultaneous equations. They are

$$\sum_1^m c_i \int_R\int \Big[\phi_i(x, y)-\lambda\int_R\int G(x, y; x', y'; 0) \times\phi_i(x', y')\rho(x', y')dx'dy'\Big] \times\phi_k(x, y)\rho(x, y)dxdy=0, \quad k=1, 2, \cdots, m \quad (8)$$

in Grammel's method and

$$\sum_1^m c_i \int\int \{L[\phi_i]+\lambda\rho\phi_i\}\phi_k dxdy=0, \quad k=1, 2, \cdots, m \quad (9)$$

in Galerkin's.

The former set of equations appear to be more formidable but actually they are more tractable and easier for calculation than the latter set. This is due to the fact that the choice of ϕ_i's in Grammel's method is always easier than in Galerkin's, because, there, no differentiations on ϕ_i's are involved. If the same expression, i.e., the same linear combination of the same ϕ_i's is used as the trial function for both cases, Grammel has shown that the upper bound calculated by his method is always lower than that calculated by Galerkin's or Ritz's method, the latter two giving exactly the same value.

In some cases we do not even need to know the explicit expression of the Green's function for carrying out the Grammel process. It is usually possible to transform the double integral into a single one

$$\int_R\int \phi_k(x, y)\rho(x, y)dxdy \times\int_R\int G(x, y; x', y'; 0)\phi_i(x', y')\rho(x', y')dx'dy' = \int_R\int \frac{M(\phi_i, \rho)\cdot M(\phi_k, \rho)}{\alpha(x, y)}dxdy,$$

where M and α can in some cases be easily recognized as known physical quantities.[9] Thus, the set of simultaneous equations may be reduced to

$$\sum_1^m c_i\Big[\int_R\int \phi_i\phi_k\rho dxdy - \lambda\int_R\int \frac{M(\phi_i, \rho)\cdot M(\phi_k, \rho)}{\alpha}dxdy\Big]=0, \quad k=1, 2, \cdots, m.$$

[9] For example, M may represent the bending moment in dynamics, then α will be the modulus of rigidity. See R. Grammel and Biezeno-Grammel, references 3 and 8.

In all the methods discussed so far, we need to set up a set of simultaneous equations and to solve them. The process is usually lengthy and difficult. Besides, the accuracy for higher modes is much worse than for the fundamental, because the values for different modes all come out of the same set of equations. No use has been made of the fact that a given eigenfunction is orthogonal to all the lower ones.

If we know approximately the first $n-1$ eigenfunctions, the knowledge can be put into immediate use for calculating the lower bound for the next higher eigenvalue by the help of the so-called maximum-minimum-property of the eigenvalues.[10] We first consider the problem with the original boundary condition $u=f(s)$ but the loosened orthogonality relations

$$\int_R\int w\cdot v_i\rho dxdy=0, \quad i=1, 2, \cdots, n-1,$$

where w is the trial function and v_i's are the first $n-1$ approximate eigenfunctions. The would-be eigenvalue for this perturbed problem will be less than the true value on account of its maximum-minimum-property. Then if we also loosen the boundary condition in a similar way as in Trefftz's method, the calculated value will be still lower than the would-be eigenvalue. The more we know about the v_i's, the better will be the calculated lower bound. However, the process is not quite so simple as briefly described above and cannot be applied for calculating upper bounds.

From the above discussion, it is to be expected that some modification of the Courant-Trefftz method formulated on the basis of the integral equation will have the merit of being both simple and flexible. In fact, such has already been furnished by Schwinger.[11] His method for calculating both upper and lower bounds proves to be very successful. It embodies Courant's idea of perturbing boundary conditions and looks like Grammel's method in using an integral equation.

Schwinger has only worked on boundary value problems, while his method can be used for the calculation of eigenvalues as well. We have done some calculations by his method on the resonant frequency of a disk-loaded cavity and have described in a previous paper[12] our results and how we proceeded with our calculations. We have explained how this method can give an upper bound in one type of calculation and a lower bound in another, one being almost as good as the other.

Despite the stringent condition that the trial functions are to satisfy the differential equation, the method can be applied to quite a large number of important physical problems. It is particularly useful for those

[10] For a proof of the property, see Courant u. Hilbert, *Methoden der Mathematischen Physik* (Verlag. Julius Springer, Berlin, 1931), Vol. I, p. 351–353.

[11] J. Schwinger, "Discontinuities in wave guides" (lecture notes prepared by D. S. Saxon).

[12] E. L. Chu and W. W. Hansen, J. App. Phys. **20**, 280–285 (1949).

problems concerning composite-type regions. A region is called composite, if it can be divided by simple surfaces into two or more regions, in each of which the problem can be exactly solved once we impose simple enough conditions on those dividing surfaces. The disk-loaded cavity, which we have considered, is one example.

To facilitate further discussions, it is desirable to give a brief summary of the method in simple concise terms. We take $G(x, y; x', y'; \lambda)$ as the Green's function of the differential equation $L[u]+\lambda\rho u=0$, such that either

$$G=0 \text{ on } \Gamma \tag{10a}$$

or

$$\partial G/\partial n=0 \text{ on } \Gamma. \tag{10b}$$

The integral equation to be solved will be either

$$\frac{\partial u}{\partial n}(s)=-\sigma u(s)=-\int_\Gamma pu\frac{\partial}{\partial n}\frac{\partial G(\lambda)}{\partial n'}ds' \tag{11a}$$

or

$$u(s)=-\frac{1}{\sigma}\frac{\partial u}{\partial n}(s)=\int_\Gamma pG(\lambda)\frac{\partial u}{\partial n'}ds' \tag{11b}$$

as the case may be. $\sigma=-(1/u)(\partial u/\partial n)$ is known as given boundary condition and is usually a piecewise continuous function of s. These equations are not exactly soluble, so we perturb the problem and solve $L[w]+\mu\rho w=0$ with boundary values related by

$$\frac{\partial w}{\partial n}(s)=-\xi w(s)=-\int_\Gamma pw\frac{\partial}{\partial n}\frac{\partial G(\mu)}{\partial n'}ds' \tag{12a}$$

or

$$w(s)=-\frac{1}{\xi}\frac{\partial w}{\partial n}(s)=\int_\Gamma pG(\mu)\frac{\partial w}{\partial n'}ds', \tag{12b}$$

where ξ may be left undefined or chosen properly. If we demand

$$\int pw\frac{\partial w}{\partial n}ds=\int_\Gamma p\frac{\partial u}{\partial n}\cdot\frac{w^2}{u}ds$$

or

$$\int pw\frac{\partial w}{\partial n}ds=\int_\Gamma p\frac{u}{\partial u/\partial n}\cdot\left(\frac{\partial w}{\partial n}\right)^2ds,$$

we obtain

$$\int_\Gamma p\frac{\partial u}{\partial n}\frac{w^2}{u}ds=-\int_\Gamma pwds\int_\Gamma pw\frac{\partial}{\partial n}\frac{\partial G(\mu)}{\partial n'}ds' \tag{13a}$$

or

$$\int_\Gamma p\frac{u}{\partial u/\partial n}\left(\frac{\partial w}{\partial n}\right)^2ds=\int_\Gamma p\frac{\partial w}{\partial n}ds\int pG(\mu)\frac{\partial w}{\partial n'}ds'. \tag{13b}$$

By assuming certain functions for $w(s)$ or $(\partial w/\partial n)(s)$, we can calculate μ from (13a) or (13b) respectively. In both types of calculation we may approach the limiting case $w=u$ and $\mu=\lambda$ as far as may be desired.

The value calculated is only an approximation to the true eigenvalue. The accuracy depends on how we perturb the boundary condition, that is, how we choose w. If we have a good approximate knowledge about u to start with, we may be able to get a very accurate value by a single trial without dealing with the simultaneous equations or solving a determinantal one.

Whether the calculated value is an upper or a lower bound depends on whether we have strengthened or weakened the supplementary conditions of the problem. Except for very simple cases, the distinction between a stringent and a weak condition is not quite obvious. Even where we are sure we have weakened the supplementary conditions, the existence of a lower bound depends on a further condition that the calculated value should be stationary with respect to any small variation in all the admissible trial functions. In the simple process described above we have not gone through such variational considerations.

Schwinger has discussed such questions on boundary value problems. He has shown that his variational expressions possess the property of being definite. Yet, further justifications are necessary in order to formulate some general rules for the calculation of eigenvalues. It is the purpose of this paper to concern ourselves with such justifications.

II. THE FUNDAMENTAL MODE

Let λ_1 and u_1 be the lowest eigenvalue and the corresponding eigenfunction for the Sturm-Liouville differential Eq. (1). u_1 will have no zeros in the region R. It is to be found by minimizing the expression

$$J[u]=D[u]+\int_\Gamma p\sigma u^2ds, \tag{14}$$

where $D[u]$ is as given in Eq. (2), under the boundary condition that $u=0$ at least on that part of Γ where u_1 is zero, and also under the normalizing condition

$$\int_R\int\rho u^2dxdy=1.$$

u should be continuous and have piecewise continuous first derivatives in R, but is otherwise arbitrary. σ is any given piecewise continuous function of coordinates on Γ, such that

$$(\partial u_1/\partial n)+\sigma u_1=0 \text{ on } \Gamma. \tag{15}$$

The expression to be minimized may be written as

$$J[u]=D[u]-\int_\Gamma p\frac{u^2}{u_1}\frac{\partial u_1}{\partial n}ds. \tag{16}$$

This is sensible because $u=0$ wherever $u_1=0$ on Γ. The expression (16) may be transformed by Green's theorem, to obtain

$$D[u]-\int_\Gamma p\frac{u^2}{u_1}\frac{\partial u_1}{\partial n}ds$$
$$=D[u]-D\left[u_1,\frac{u^2}{u_1}\right]-\int_R\int\frac{u^2}{u_1}L[u_1]dxdy,$$

where

$$D[u,v]=\int_R\int p(u_xv_x+u_yv_y)dxdy+\int_R\int quvdxdy.$$

Thus

$$D[u]-\int_\Gamma p\frac{u^2}{u_1}\frac{\partial u_1}{\partial n}ds=\int_R\int\frac{p}{u_1^2}(u_1\nabla u-u\nabla u_1)^2dxdy$$
$$-\int_R\int\frac{u^2}{u_1}L[u_1]dxdy.$$

Since

$$L[u_1]+\lambda_1\rho u_1=0,$$

we get

$$D[u]-\int_\Gamma p\frac{u^2}{u_1}\frac{\partial u_1}{\partial n}ds$$
$$=\int_R\int\frac{p}{u_1^2}(u_1\nabla u-u\nabla u_1)^2dxdy$$
$$+\lambda_1\int_R\int\rho u^2dxdy. \quad (17)$$

In carrying out the above transformation, it is tacitly assumed that u has first derivatives everywhere in R. If not, we may consider R as being composed of several regions in each of which u is everywhere differentiable. Equation (17) brings out the variational principle explicitly in view. It states that λ_1 is the minimum value of $J[u]$ for any admissible u that is properly normalized, because

$$\int_R\int\frac{p}{u_1^2}(u_1\nabla u-u\nabla u_1)^2dxdy\geqq 0.$$

Now we restrict ourselves to use only those trial functions which satisfy the differential equation $L[u]+\lambda\rho u=0$. From (17) and

$$D[u]=-\int_R\int uL[u]dxdy+\int_\Gamma pu\frac{\partial u}{\partial n}ds$$
$$=\lambda\int_R\int\rho u^2dxdy+\int_\Gamma pu\frac{\partial u}{\partial n}ds,$$

we obtain

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds-\int_\Gamma p\frac{u^2}{u_1}\frac{\partial u_1}{\partial n}ds$$
$$=\int_R\int\frac{p}{u_1^2}(u_1\nabla u-u\nabla u_1)^2dxdy$$
$$-(\lambda-\lambda_1)\int_R\int\rho u^2dxdy. \quad (18)$$

If we demand

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds-\int_\Gamma p\frac{u^2}{u_1}\frac{\partial u_1}{\partial n}ds\leqq 0,$$

then

$$\lambda\geqq\lambda_1.$$

Thus we formulate a simple rule for calculating an upper bound for λ_1 as follows:

Choose a function $u(s)$ such that $u(s)=0$ on Γ wherever $\sigma=\infty$. Derive $\partial u/\partial n$ from $u(s)$ such that $u(x,y)$ satisfies $L[u]+\lambda\rho u=0$. Solve the equation

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds=-\int_\Gamma p\sigma u^2ds \quad (19)$$

for λ. The value obtained is always an upper bound.

The method of solution is as follows:

Since

$$u(x,y)=-\int_\Gamma pu\frac{\partial G}{\partial n'}ds',$$

the derived expression for $\partial u/\partial n$ is

$$\frac{\partial u}{\partial n}=-\int_\Gamma pu\frac{\partial}{\partial n}\frac{\partial G}{\partial n'}ds',$$

where $G(x,y;x',y';\lambda)$ is the Green's function of the equation $L[u]+\lambda\rho u=0$, such that $G=0$ on Γ.[13] Thus

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds=-\int_\Gamma puds\int_\Gamma pu\frac{\partial}{\partial n}\frac{\partial G}{\partial n'}ds'$$
$$=\int_\Gamma\int_\Gamma p(s)u(s)K(s,s';\lambda)p(s')u(s')dsds', \quad (20)$$

with

$$K(s,s';\lambda)=-\left[\frac{\partial}{\partial n}\frac{\partial}{\partial n'}G(x,y;x',y';\lambda)\right]_{\substack{x,y\text{ on }\Gamma\\ x',y'\text{ on }\Gamma,}}$$

and (19) becomes

$$\int_\Gamma\int_\Gamma p(s)u(s)K(s,s';\lambda)p(s')u(s')dsds'=-\int_\Gamma p\sigma u^2ds.$$

In some cases it is easier to find $K(s,s';\lambda)$ than it is to find G and then differentiate it. From this equation λ can be determined when $u(s)$ is chosen. λ depends, of course, on $u(s)$ but the above rule implies that this equation can never be satisfied with $\lambda<\lambda_1$ for any admissible $u(s)$.

The function $u(s)$ should be a good approximation to

[13] It is expedient to have $G=0$ only on that part of Γ, say Γ', where the boundary condition is perturbed and $\partial G/\partial n+\sigma G=0$ on the rest of Γ. Then G is appropriate on $\Gamma-\Gamma'$ and $u(x,y)=-\int_{\Gamma'}pu(\partial G/\partial n')ds'$.

$u_1(s)$ and should be so chosen that the integration of $\int_\Gamma pu(\partial u/\partial n)ds$ will be easy. With a good choice of $u(s)$, which is usually practicable, the derived function $u(x, y)$ will be a very good approximation to $u_1(x, y)$, so the calculated value of λ can be very accurate. The Green's function G or the kernel K is usually expressed in the form of an infinite series; we need to sum it before we can solve the Eq. (19). The algebraic process of summing a series, though sometimes difficult, is much simpler than the solution of a high order determinant.

To explain how to obtain a lower bound for the eigenvalue, we interchange u with u_1 and λ with λ_1 in Eq. (18), and thus obtain an equation which is the mate of (18):

$$\int_\Gamma pu_1\frac{\partial u_1}{\partial n}ds-\int_\Gamma p\frac{u_1^2}{u}\frac{\partial u}{\partial n}ds$$

$$=\int_R\int\frac{p}{u^2}(u\nabla u_1-u_1\nabla u)^2dxdy$$

$$-(\lambda_1-\lambda)\int\int\rho u_1^2dxdy. \quad (21)$$

Here we must have the condition that $u_1=0$ on Γ wherever $u=0$ on Γ. In other words, u can be zero on Γ only where u_1 is known to be zero. Furthermore, u can have no zeros inside R. The latter restriction implies that u must be the solution for the fundamental mode of the modified problem.

If we demand

$$\int_\Gamma pu_1\frac{\partial u_1}{\partial n}ds-\int_\Gamma p\frac{u_1^2}{u}\frac{\partial u}{\partial n}ds\leqq 0, \quad (22)$$

we should have $\lambda\leqq\lambda_1$. But this is not as practicable as the upper-bound case, because u_1 and $u_1(\partial u_1/\partial n)$ are not supposed to be known simultaneously on Γ except where $u_1=0$. The boundary condition is given by $(1/u_1)(\partial u_1/\partial n)$, not by both u_1 and $\partial u_1/\partial n$. Thus we have to demand $(1/u)(\partial u/\partial n)\geqq(1/u_1)(\partial u_1/\partial n)$ everywhere on that part of Γ where $u_1\neq 0$ in order that the condition (22) will be certainly satisfied. If we are going to perturb the boundary condition slightly, the modified problem may not be exactly soluble; if we perturb more, the calculated value may not be accurate enough.

One important case, however, can be dealt with immediately. It occurs most often in practical problems especially those concerning mechanical and electromagnetic oscillations. In this case, the boundary condition is given as $u_1(\partial u_1/\partial n)=0$ everywhere on Γ; so if we choose u such that $\partial u/\partial n=0$ wherever $u_1\neq 0$ (i.e., $\partial u_1/\partial n=0$) on Γ, we will have $\lambda\leqq\lambda_1$.

In general, it is possible to formulate a rule similar to that for the upper-bound case. The rule is as follows:

Choose a function $(\partial u/\partial n)(s)$ such that $\partial u/\partial n=0$ on Γ wherever $\sigma=0$. Derive $u(s)$ from $\partial u/\partial n$ such that $u(x, y)$ satisfies $L[u]+\lambda\rho u=0$. Solve the equation

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds=-\int_\Gamma\frac{p}{\sigma}\left(\frac{\partial u}{\partial n}\right)^2ds \quad (23)$$

for λ. The value obtained is always a lower bound.

It must be understood here that $\partial u/\partial n$ should be reasonably close to $\partial u_1/\partial n$, otherwise we might get a solution for a higher mode.

The solution is obtained in a similar way as in the upper-bound case.

Here $u(x, y)=\int_\Gamma pG(\partial u/\partial n')ds'$, in which $G(x, y; x', y'; \lambda)$ is the Green's function of $L[u]+\lambda\rho u=0$ such that $(\partial G/\partial n)=0$ on Γ.[14] Thus

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds=\int\int_\Gamma p(s)\frac{\partial u}{\partial n}G(s, s'; \lambda)p(s')\frac{\partial u}{\partial n'}dsds'. \quad (24)$$

The above rule implies that the equation

$$\int_\Gamma\int_\Gamma p(s)\frac{\partial u}{\partial n}G(s, s'; \lambda)p(s')\frac{\partial u}{\partial n'}dsds'=-\int_\Gamma\frac{p}{\sigma}\left(\frac{\partial u}{\partial n}\right)^2ds$$

can never be satisfied with $\lambda>\lambda_1$ for any admissible trial function.

The proof of the rule, however, is not so easy as for the upper-bound case. We are not going to be concerned with it here, since it is but a special case of that for the higher modes, for which a general proof is given below.

So far, we have only considered the perturbation of boundary conditions. Similar results will be obtained by perturbation of the conditions of continuity. Discussion of the latter case will also be reserved for the higher modes.

III. HIGHER MODES

1. Upper Bound

We consider the lth mode. Let λ_l and u_l be the lth eigenvalue and the corresponding eigenfunction. u_l has zeros (nodes or nodal lines or nodal surfaces) inside R; they divide R into no more than l sub-regions.[15] It is to be found by minimizing the same expression (14) as for the fundamental mode. The supplementary conditions as to boundary values, the normalizing condition and the continuity requirements are also the same, but there are $l-1$ additional orthogonality conditions. They are:

$$\int_R\int\rho uu_i dxdy=0, \quad i=1, 2, \cdots, l-1. \quad (25)$$

For the case of a higher mode, the variational expression cannot be transformed into a form as represented by Eq. (17), except when u and u_l have exactly the same zeros inside R. In general, such transformation

[14] See reference 13. If $(\partial G/\partial n)=0$ on Γ' and G is appropriate on $\Gamma-\Gamma'$, then $u(x, y)=\int_{\Gamma'}pG(\partial u/\partial n')ds'$.

[15] Courant u. Hilbert, reference 10, S. 393.

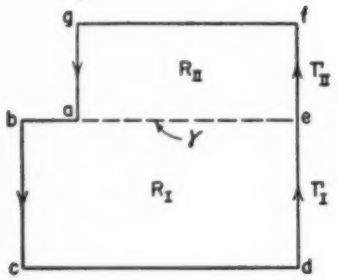


FIG. 1. A composite region R is shown divided into two simple regions R_I and R_{II} by the surface γ. The open boundary *abcde* is denoted by Γ_I and *efga* by Γ_{II}.

cannot be justified. We may, of course, cut out those singularities and then transform the expression, but the result so obtained does not appear to have an explicit extremal property.

It is easier to proceed directly from the variational principle. With u satisfying all the appropriate supplementary conditions mentioned above, the principle assures us that the expression has a minimum value equal to λ_1. The minimum is reached when and only when $u=u_1$. Thus we have, for any admissible u,

$$D[u]+\int_\Gamma p\sigma u^2 ds \geqq \lambda_1 \int_R\int \rho u^2 dxdy. \qquad (26)$$

If u satisfies the differential equation $L[u]+\lambda\rho u=0$, the inequality (26) may be transformed to

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds+\int_\Gamma p\sigma u^2 ds \geqq (\lambda_1-\lambda)\int_R\int \rho u^2 dxdy. \qquad (27)$$

If we demand, as before,

$$\int_\Gamma pu\frac{\partial u}{\partial n}ds+\int_\Gamma p\sigma u^2 ds \leqq 0,$$

then

$$\lambda \geqq \lambda_1.$$

The rule for calculating an upper bound for a higher eigenvalue is just the same as for the lowest except that the trial functions must be orthogonal to all the lower eigenfunctions. If we really need to know the exact form of these latter functions, the method would not be of any practical use. For problems which cannot be solved exactly, we have, at most, only an approximate knowledge about these lower eigenfunctions. If we use approximate functions instead of true ones, the calculated value will be somewhat lower than it would be according to the maximum-minimum-property of eigenvalues. Thus we are not sure whether it is really an upper bound. With the property that u satisfies the differential equation, as it must in this method, this kind of difficulty can easily be removed. We will show how this can be done in a later section.

Now we come to the question of the perturbation of continuity conditions. Let us divide the region R into R_I and R_{II} by γ, as shown in Fig. 1. Let Γ_I and Γ_{II} denote those parts of Γ which belong to R_I and R_{II}, respectively. We choose a function $u_I(s)$ on $\Gamma_I+\gamma$ and $u_{II}(s)$ on $\Gamma_{II}+\gamma$ with $u_I=u_{II}$ on γ, and then derive the trial functions $u_I(x, y)$ and $u_{II}(x, y)$ for the two regions such that they both satisfy the same differential equation.

Thus, u is continuous everywhere in R but the normal component of ∇u is in general discontinuous on γ. u is to satisfy all those restrictions imposed on the trial function for a higher mode for the case of perturbation of boundary conditions, but has no restriction on γ except that it should be continuous there. From the inequality (26), we have instead of (27)

$$\int_{\Gamma_I+\gamma} pu_I\frac{\partial u_I}{\partial n}ds+\int_{\Gamma_{II}+\gamma} pu_{II}\frac{\partial u_{II}}{\partial n}ds + \int_{\Gamma_I+\Gamma_{II}} p\sigma u^2 ds \geqq (\lambda_1-\lambda)\int_R\int \rho u^2 dxdy. \qquad (28)$$

The integrals

$$\int_{\Gamma_I+\gamma} pu_I\frac{\partial u_I}{\partial n}ds \quad \text{and} \quad \int_{\Gamma_{II}+\gamma} pu_{II}(\partial u_{II}/\partial n)ds$$

can be expressed as similar functions of $u_I(s)$ and $u_{II}(s)$ as shown by Eq. (20). We obtain an upper bound for λ_1 by solving the following equation

$$\int_{\Gamma_I+\gamma} pu\frac{\partial u}{\partial n}ds+\int_{\Gamma_{II}+\gamma} pu\frac{\partial u}{\partial n}ds=-\int_\Gamma p\sigma u^2 ds.$$

2. Lower Bound

We perturb the problem by replacing the boundary condition $\partial u_1/\partial n+\sigma u_1=0$ with $\partial w/\partial n+\xi w=0$ such that w does not have to be zero on that part of Γ where $u_1=0$ and $\xi(s)$ is finite wherever $\sigma(s)$ is finite. $\xi(s)$ is a piecewise continuous function and when chosen is to be considered as fixed. The modified problem is to minimize the expression

$$J_1[v]=D[v]+\int_\Gamma p\xi v^2 ds \qquad (29)$$

under the same orthogonality, normalizing and continuity conditions as those imposed on u for minimizing the expression

$$J[u]=D[u]+\int_\Gamma p\sigma u^2 ds$$

of the original problem. Only the boundary conditions are different. Here $v(s)=0$ wherever $\xi(s)$ is infinite, i.e., $w(s)=0$ and $v(s)$ may not be zero where $u_1(s)=0$.

Let us denote the set of trial functions for the original problem by u and that for the modified problem by v. It is obvious that every function of u is contained in v, so u_1 is admissible as a trial function for the expression (29). Let w be that function of the set v which minimizes $J_1[v]$ and μ its minimum value, i.e., $J_1[v]_{\min}=J_1[w]$

$=\mu$. According to the usual variational principle w must satisfy the differential equation $L[w]+\mu\rho w=0$ besides the boundary condition $\partial w/\partial n+\xi w=0$.

Since $J_1[u_l]\geqq J_1[w]$, we obtain easily

$$\int pu_l\frac{\partial u_l}{\partial n}ds+\int_\Gamma p\xi u_l^2 ds\geqq\mu\int_R \rho w^2 dxdy-\lambda_l\int \rho u_l^2 dxdy,$$

i.e.,

$$\int_\Gamma pu_l\frac{\partial u_l}{\partial n}ds-\int_\Gamma p\frac{u_l^2}{w}\frac{\partial w}{\partial n}ds\geqq\mu-\lambda_l. \tag{30}$$

Thus, if

$$\int_\Gamma pu_l\frac{\partial u_l}{\partial n}ds-\int_\Gamma p\frac{u_l^2}{w}\frac{\partial w}{\partial n}ds\leqq 0,\quad \mu\leqq\lambda_l \tag{31a}$$

and conversely if

$$\mu\geqq\lambda_l,\quad \int_\Gamma pu_l\frac{\partial u_l}{\partial n}ds-\int_\Gamma p\frac{u_l^2}{w}\frac{\partial w}{\partial n}ds\geqq 0. \tag{31b}$$

The inequality (30) is the generalization of (21) for higher modes.

As mentioned before in connection with the fundamental mode, the condition (31a) is not quite practical because u_l and $u_l(\partial u_l/\partial n)$ are not known simultaneously on Γ except for certain special cases. In fact, (31a) is not a necessary condition for $\mu\leqq\lambda_l$ and there may exist other conditions which can better serve the purpose. To find one of more general usefulness, it is natural to think of replacing u_l in (31a) by w. One way of doing this is to demand instead of (31a)

$$\int_\Gamma p\xi w^2 ds-\int_\Gamma p\sigma w^2 ds=0,$$

but this necessitates that $w(s)=0$ wherever σ is infinite and is only soluble for $\mu\geqq\lambda_l$.

Alternatively we may demand the following condition

$$\int_\Gamma p\xi w^2 ds-\int_\Gamma\frac{p}{\sigma}\left(\frac{\partial w}{\partial n}\right)^2 ds=0,$$

i.e.,

$$\int\frac{p}{\xi}\left(\frac{\partial w}{\partial n}\right)^2 ds-\int_\Gamma\frac{p}{\sigma}\left(\frac{\partial w}{\partial n}\right)^2 ds=0 \tag{32a}$$

with $\partial w/\partial n=0$ on Γ wherever $\sigma=0$. Since w satisfies the differential equation $L[w]+\mu\rho w=0$, the above condition may be written as

$$-\int_\Gamma p\frac{\partial w}{\partial n}ds\int p\frac{\partial w}{\partial n'}G(s,s';\mu)ds'-\int_\Gamma\frac{p}{\sigma}\left(\frac{\partial w}{\partial n}\right)^2 ds=0. \tag{32b}$$

Here $G(x,y;x',y';\mu)$ is the Green's function of $L[w]+\mu\rho w=0$ such that $\partial G/\partial n=0$ on Γ. It will be shown immediately that the condition (32a), i.e., (32b) is indeed sufficient to make $\mu\leqq\lambda_l$ for any admissible trial function used for $\partial w/\partial n$.

To prove this, we first note that if $\xi=\sigma$ except where $\xi=0$ but $\sigma>0$, (31a) is certainly satisfied and (32a) becomes an identity. For any such ξ we have a unique solution with $\mu\leqq\lambda_l$. In other words, with such values of μ we can always find a function $\partial w/\partial n$ which satisfies (32b). However, we need to prove the converse statement, i.e., if $\mu>\lambda_l$ (32b) can never be satisfied provided μ is reasonably close to λ_l.

Next we consider w to be in a small neighborhood of u_l and take $\mu=\lambda_l$, i.e., $L[w]+\lambda_l\rho w=0$. Let us denote $w=u_l+\delta w$, $\xi=\sigma+\epsilon$ and

$$\delta J(u_l)=\int_\Gamma p\xi u_l^2 ds-\int p\sigma u_l^2 ds,$$

δw and ϵ being functions of small magnitudes. Then from (31b) we have

$$\delta J(u_l)=\int_\Gamma p\epsilon u_l^2 ds\geqq 0.$$

As is evident from the extremal property of $\delta J(u_l)$, it should be possible to express $\delta J(u_l)$ in terms of second-order small quantities. This can be easily shown as follows:

$$\delta J(u_l)=\int_\Gamma pu_l\frac{\partial u_l}{\partial n}ds-\int_\Gamma p\frac{u_l^2}{w}\frac{\partial w}{\partial n}ds$$

$$=\int_\Gamma p\delta w\frac{\partial\delta w}{\partial n}ds-\int_\Gamma p\frac{(\delta w)^2}{w}\frac{\partial w}{\partial n}ds.$$

Then we consider the corresponding variation for the expression (32a)

$$\delta J(w)=\int_\Gamma\frac{p}{\xi}\left(\frac{\partial w}{\partial n}\right)^2 ds-\int_\Gamma\frac{p}{\sigma}\left(\frac{\partial w}{\partial n}\right)^2 ds$$

$$=-\int p\delta w\frac{\partial\delta w}{\partial n}ds-\int_\Gamma\frac{p}{\sigma}\left(\frac{\partial\delta w}{\partial n}\right)^2 ds,$$

or

$$\delta J(w)=-\int_\Gamma p\frac{\epsilon}{\sigma(\sigma+\epsilon)}\left(\frac{\partial w}{\partial n}\right)^2 ds$$

$$=-\int_\Gamma p\epsilon u_l^2\frac{1+[(\partial\delta w/\partial n)/(\partial u_l/\partial n)]^2}{1+\epsilon/\sigma}ds.$$

As long as small variations are considered, $[(\partial\delta w/\partial n)/(\partial u_l/\partial n)]^2\ll 1$ and $\epsilon/\sigma\ll 1$ except in the small neighborhoods of those zeros of $u_l(s)$ where $\partial u_l/\partial n$ is also zero. Let us denote such neighborhoods by η, then

$$\delta J(w)=\delta J_{\Gamma-\eta}+\delta J_\eta$$

where

$$\delta J_\eta = -\int_\eta p\delta w \frac{\partial \delta w}{\partial n} ds - \int_\eta \frac{p}{\sigma}\left(\frac{\partial \delta w}{\partial n}\right)^2 ds$$

and $\delta J_{\Gamma-\eta}$ is a similar expression. As $\partial\delta w/\partial n \to 0$, $\eta \to 0$, $|\delta J_\eta|$ is small in comparison with $|\delta J_{\Gamma-\eta}|$ and $\delta J(w) \to \delta J_{\Gamma-\eta} \to -\int_\Gamma p\epsilon u_l^2 ds$. Since $\int_\Gamma p\epsilon u_l^2 ds \geqq 0$, we must have $\delta J(w) \leqq 0$ at least for small enough variations.

Thus far it has been assumed that ξ is finite wherever σ is finite. We will now remove this restriction.

Let w' be a function which satisfies $L[w'] + \lambda_l \rho w' = 0$ and the given orthogonality conditions. w' is to be in a small neighborhood of u_l; $\partial w'/\partial n = \partial u_l/\partial n + \partial \delta w'/\partial n$, $w' = u_l + \delta w'$. $\partial w'/\partial n = 0$ on Γ wherever $\sigma = 0$. $\xi' = -(1/w')(\partial w'/\partial n)$ may be infinite where σ is finite. By small variations w' may be changed into a function, say w, which satisfies all the conditions imposed on w' and the restriction that $\xi = -(1/w)(\partial w/\partial n)$ is finite where σ is finite. We denote

$$\partial w'/\partial n = \partial w/\partial n + \partial \delta w''/\partial n \qquad (\partial \delta w''/\partial n = \partial \delta w'/\partial n - \partial \delta w/\partial n),$$

$$w' = w + \delta w'' \qquad (\delta w'' = \delta w' - \delta w).$$

As long as $\partial \delta w''/\partial n$ is small, the corresponding zeros of $w(s)$ and $w'(s)$ are near together. Since w may be any suitable function, $\partial \delta w''/\partial n$ may be chosen to be zero on Γ except in those small neighborhoods of the zeros of $w(s)$ and $w'(s)$, say η, where σ and ξ are finite and $(\partial \delta w''/\partial n)$ is to be of suitable form in order that w may satisfy all prescribed conditions.

By simple calculation $\delta J(w')$ is found to be

$$\delta J(w') = \int_\Gamma \frac{p}{\xi}\left(\frac{\partial w'}{\partial n}\right)^2 ds - \int_\Gamma \frac{p}{\sigma}\left(\frac{\partial w'}{\partial n}\right)^2 ds = \delta J(w) + \delta J_\eta',$$

where

$$\delta J_\eta' = -\int_\eta p\delta w'' \frac{\partial \delta w''}{\partial n} ds - \int_\eta \frac{p}{\sigma}\left(\frac{\partial \delta w''}{\partial n}\right)^2 ds + 2\int_\eta \frac{p}{\sigma} \epsilon w \frac{\partial \delta w''}{\partial n} ds.$$

As $\partial \delta w'/\partial n$ and $\partial \delta w/\partial n$ approach zero both $\partial \delta w''/\partial n$ and η approach zero. If $\partial \delta w''/\partial n$ is sufficiently small, η will be so small that $|\delta J_\eta'| < |\delta J(w)|$. Therefore $\delta J(w') \leqq 0$ at least for small enough variations.

Since $\delta J(w')$ can also be expressed in terms of second-order small quantities as

$$\delta J(w') = -\int_\Gamma pw' \frac{\partial w'}{\partial n} ds - \int_\Gamma \frac{p}{\sigma}\left(\frac{\partial w'}{\partial n}\right)^2 ds = -\int_\Gamma p\delta w' \frac{\partial \delta w'}{\partial n} ds - \int_\Gamma \frac{p}{\sigma}\left(\frac{\partial \delta w'}{\partial n}\right)^2 ds$$

and $\delta w'$ is linearly related to $\partial \delta w'/\partial n$, so if $\delta J(w') \leqq 0$ for small $\delta w'$ the same inequality will be true for any large $\Delta w'$ which is a constant multiple of $\delta w'$. Since $\delta w'$ may be any admissible variation, so $\Delta w'$ may be any admissible function. Writing $w = u_l + \Delta w'$ we obtain

$$-\int_\Gamma pw \frac{\partial w}{\partial n} ds - \int_\Gamma \frac{p}{\sigma}\left(\frac{\partial w}{\partial n}\right)^2 ds \leqq 0. \tag{33}$$

Here w is no longer restricted to a small neighborhood of u_l and $\xi = -(1/w)(\partial w/\partial n)$ may be left undefined provided that $\partial w/\partial n = 0$ wherever $\sigma = 0$, $L[w] + \lambda_l \rho w = 0$ and w satisfies all those orthogonality conditions imposed on the set of functions u or v.

We then calculate

$$\frac{\partial}{\partial \mu}\int_\Gamma \frac{p}{\xi}\left(\frac{\partial w}{\partial n}\right)^2 ds = -\frac{\partial}{\partial \mu}\int_\Gamma pw \frac{\partial w}{\partial n} ds$$

for fixed $\partial w/\partial n$ with $L[w] + \mu\rho w = 0$.

$$-2\frac{\partial}{\partial \mu}\int_\Gamma pw \frac{\partial w}{\partial n} ds = -2\int_\Gamma p \frac{\partial w}{\partial \mu}\frac{\partial w}{\partial n} ds$$

$$= -2D\left[w, \frac{\partial w}{\partial \mu}\right] + 2\mu\int_R\int \rho w \frac{\partial w}{\partial \mu} dxdy$$

$$= -\frac{\partial}{\partial \mu}\left\{D[w] - \mu\int_R\int \rho w^2 dxdy\right\} - \int_R\int \rho w^2 dxdy$$

$$= -\frac{\partial}{\partial \mu}\int_\Gamma pw \frac{\partial w}{\partial n} ds - \int_R\int \rho w^2 dxdy.$$

Thus

$$-\frac{\partial}{\partial \mu}\int_\Gamma pw \frac{\partial w}{\partial n} ds = -\int_R\int \rho w^2 dxdy,$$

so

$$\frac{\partial}{\partial \mu}\int_\Gamma \frac{p}{\xi}\left(\frac{\partial w}{\partial n}\right)^2 ds < 0. \tag{34}$$

This inequality is valid everywhere except at certain singular points at which the expression

$$\int_\Gamma p \frac{\partial w}{\partial n} ds \int_\Gamma p \frac{\partial w}{\partial n'} G(s, s'; \mu) ds'$$

goes to infinity and changes sign abruptly.

For $\mu = \lambda_l$ the expression

$$-\int_\Gamma pw \frac{\partial w}{\partial n} ds - \int_\Gamma \frac{p}{\sigma}\left(\frac{\partial w}{\partial n}\right)^2 ds$$

is always less than or equal to zero and its value for any fixed $(\partial w/\partial n)$ can only be increased by decreasing μ, assuming μ to be varied over a range which does not contain the singular points. Therefore the condition (32)

can only be satisfied with $\mu \leqq \lambda_l$ no matter what function is chosen for $\partial w/\partial n$ as long as it is an admissible one and μ is less than the lowest singular point lying above λ_l.

We have thus proved our rule for the calculation of a lower bound for the lth eigenvalue. The rule is just the same as that stated in connection with the fundamental mode except that we have, now, $l-1$ additional orthogonality conditions. In passing, it may be noted that we have only used the orthogonality conditions at $\mu=\lambda_l$ in the above proof.

Regarding the question about the perturbation of continuity conditions, we consider the same region R as shown in Fig. 1, and consider the same problem, that is to minimize the expression (14) under the usual appropriate conditions. This is equivalent to a problem of finding a function u which minimizes

$$J[u]=D_I[u]+\int_{\Gamma_I} p\sigma u^2 ds - \int_{\gamma_I} pu(\partial u/\partial n)ds + \text{similar terms for } II \quad (35)$$

under the additional restriction that both u and $\partial u/n$ are continuous on the common boundary γ, i.e.,

$$\int_{\gamma_I+\gamma_{II}} pu(\partial u/\partial n)ds \equiv 0.$$

Here the subscripts I and II refer to regions R_I and R_{II} respectively.

We then perturb the problem such that only the normal derivative of the function should be continuous, not necessarily the function itself and the two surface integrals over γ should add up to zero. Let v denote the new set of trial functions; v certainly contains the set u. Therefore, some functions in v will have

$$J[v] \leqq J[u]_{\min} = J[u_l] = \lambda_l.$$

Let w be a sub-set of functions contained in v such that every function of w has a $J[w]$ which is a minimum among all functions of v which have the same normal derivative $\partial w/\partial n$ on γ. u_l is contained in w.

If the set v has functions with $J[v]<\lambda_l$, w must also have functions with $J[w]<\lambda_l$. This is obvious because there exists a function w corresponding to any such function v such that $\partial w/\partial n=\partial v/\partial n$ on γ and $J[w] \leqq J[v]$.

Since $J[u_l]$ is stationary with respect to arbitrary variations in the normal derivative on γ while other $J[w]$ are not, we must have either $J[w] \leqq J[u_l]$ or $J[w] \geqq J[u_l]$. Otherwise the second variation of $J[w]$ should vanish identically for $w=u_l$, but this can easily be shown to be impossible. Moreover, since it is always possible to find a function w which satisfies $J[w] \leqq J[u_l,]$ e.g., $\partial w/\partial n=0$ on γ is such a function, therefore we must have

$$J[w] \leqq J[u_l].$$

The case where $J[w]$ is close to a higher eigenvalue is to be excluded.

We denote $J[w]=\mu$. It follows directly from the usual variational treatment that we must have

$$L[w]+\mu\rho w=0 \text{ in } R, \quad \partial w/\partial n+\sigma w=0 \text{ on } \Gamma.$$

Under these conditions

$$w_I=\int_\gamma p\frac{\partial w_I}{\partial n'}G_I(s, s'; \mu)ds',$$

$$w_{II}=\int_\gamma p\frac{\partial w_{II}}{\partial n'}G_{II}(s, s'; \mu)ds',$$

$$\left(\frac{\partial w_I}{\partial n}=-\frac{\partial w_{II}}{\partial n} \text{ on } \gamma\right),$$

where $G_I(x, y; x', y'; \mu)$ and $G_{II}(x, y; x', y'; \mu)$ are Green's functions of the same equation $L[u]+\mu\rho u=0$, such that G satisfies the appropriate boundary conditions on Γ and $\partial G_I/\partial n=\partial G_{II}/\partial n=0$ on γ. Both G_I and G_{II} exist because we have postulated that the problem can be solved exactly for separate regions once we impose boundary conditions on γ.

Hence we may conclude that the condition

$$0=\int_{\gamma_I+\gamma_{II}} pw\frac{\partial w}{\partial n}ds = \int_\gamma p\frac{\partial w_I}{\partial n}ds\int_\gamma p\frac{\partial w_I}{\partial n'}G_I(s, s'; \mu)ds' + \int_\gamma p\frac{\partial w_{II}}{\partial n}ds\int_\gamma p\frac{\partial w_{II}}{\partial n'}G_{II}(s, s'; \mu)ds' \quad (36)$$

can never be satisfied with $\mu>\lambda_l$ for any admissible trial function used for $\partial w/\partial n$ on γ, provided μ is reasonably close to λ_l.

We may further point out that the two processes of perturbation, that of perturbing boundary conditions and of perturbing continuity conditions, may be carried out together for the same problem and the above proofs can easily be extended to such cases.

IV. SERIES EXPANSION AND ORTHOGONALITY CONDITIONS

We will show in this section how we can avoid the difficulty of having no exact knowledge about the lower eigenfunctions. We consider the case of perturbation of continuity conditions, the other case, namely, the perturbation of boundary conditions is somewhat simpler. The variational expression for the upper-bound case has properties closely connected to those of the expression for calculating the lower bound, so we will discuss them together.

We consider the lth mode for the same problem as before. Either $w(s)$ or $\partial w/\partial n$ is to be continuous on γ (Fig. 1) and $\partial w/\partial n+\sigma w(s)=0$ on Γ. In either case, we

may expand $w_I(s)$ and $w_{II}(s)$ into series in terms of different sets of orthogonal functions. This implies that we may expand G_I, G_{II} or $\partial G_I/\partial n$, $\partial G_{II}/\partial n$ in terms of their respective sets of eigenfunctions.

Thus we write

$$\begin{aligned} w_I(s)&=\sum_{j=1}^{\infty} b_{I,j}\phi_{I,j}(s) \text{ on } \gamma \\ w_{II}(s)&=\sum_{j=1}^{\infty} b_{II,j}\phi_{II,j}(s) \text{ on } \gamma, \end{aligned} \tag{37}$$

where b_j's are functions of λ but not of position, and $\phi_j(s)$ are the orthonormal coordinate functions such that

$$\int p\phi_{I,i}\phi_{I,j}ds=\delta_{ij}$$

and

$$\int_{\gamma} p\phi_{II,i}\phi_{II,j}ds=\delta_{ij}. \tag{38}$$

The existence of such functions depends, of course, on the assumption that exact solutions for the modified problem are known for R_I and R_{II} separately. In general, $\phi_{I,i}$ are not orthogonal to $\phi_{II,i}$.

We denote as before the first $(l-1)$ eigenfunctions of the original problem by u_i $(i=1, 2, \cdots, l-1)$. $u_i(s)$ may also be expanded in terms of ϕ_I's or ϕ_{II}'s. Thus

$$\begin{aligned} u_i(s)&=\sum_{j=1}^{\infty} \epsilon_{I,ij}\phi_{I,j}(s) \text{ on } \gamma \\ &=\sum_{j=1}^{\infty} \epsilon_{II,ij}\phi_{II,j}(s) \text{ on } \gamma \\ i&=1, 2, \cdots, l-1. \end{aligned} \tag{39}$$

Instead of (37) $w(s)$ may be expressed in a somewhat different form by using the $(l-1)$ functions $u_i(s)$ to replace any $(l-1)$ coordinate functions ϕ_j. Let $l-1=2v$ or $2v-1$ according as l is odd or even. We replace the first $(l-1)\phi_j$'s by u_i's and write $w(s)$ as

$$\begin{aligned} w_I(s)&=\sum_{i=1}^{l-1} c_iu_i(s)+\sum_{j=v+1}^{\infty} c_{I,j}\phi_{I,j}(s) \\ w_{II}(s)&=\sum_{i=1}^{l-1} c_iu_i(s)+\sum_{j=v \text{ or } v+1}^{\infty} c_{II,j}\phi_{II,j}(s). \end{aligned} \tag{40}$$

Substituting the expression (39) for $u_i(s)$ into (40) and comparing with (37), we obtain

$$\begin{aligned} b_{I,j}&=\sum_{i=1}^{l-1} c_i\epsilon_{I,ij}, \quad j=1, 2, \cdots, v; \\ b_{II,j}&=\sum_{i=1}^{l-1} c_i\epsilon_{II,ij}, \quad j=1, 2, \cdots, v-1 \text{ or } v. \end{aligned} \tag{41}$$

If the determinant $|\epsilon_{ij}|\neq 0$, the above $l-1$ equations will just be enough for the determination of $l-1$ c_i's. In case all b_j's vanish, c_i's can only be zero. If $|\epsilon_{ij}|=0$, we may write $w(s)$ as

$$\begin{aligned} w_I(s)&=\sum_{i=1}^{l-1} c_iu_i(s)+c_Nu_N(s)+\sum_{j=v+1}^{\infty} c_{I,j}\phi_{I,j}(s), \\ w_{II}(s)&=\sum_{i=1}^{l-1} c_iu_i(s)+c_Nu_N(s)+\sum_{j=v \text{ or } v+1}^{\infty} c_{II,j}\phi_{II,j}(s), \end{aligned}$$

with N being any integer $\geqq l+1$, such that the rank of the matrix $(\epsilon_{ij}, \epsilon_{N,j})$ is $l-1$. Then we have in place of (41)

$$\begin{aligned} b_{I,j}&=\sum_{i=1}^{l-1} c_i\epsilon_{I,ij}+c_N\epsilon_{I,Nj}, \qquad j=1, 2, \cdots, v, \\ b_{II,j}&=\sum_{i=1}^{l-1} c_i\epsilon_{II,ij}+c_N\epsilon_{II,Nj}, \quad j=1, 2, \cdots, v-1 \text{ or } v. \end{aligned}$$

By assigning a certain value, zero or non-zero, to one suitable coefficient c_i we will be able to determine the other c_i's and c_N uniquely. The rest of the discussion will be seen to be essentially the same for the two different cases. Therefore, we may assume without loss of generality that $|\epsilon_{ij}|\neq 0$ and $w(s)$ can be expressed in the form of (40).

Similarly, we can also express $\partial w/\partial n$ in terms of ϕ_j's and $u_i(s)$ as

$$\begin{aligned} \frac{\partial w_I}{\partial n}&=-\sum_{i=1}^{l-1} c_i\chi_{I,i}u_i(s)-\sum_{j=v+1}^{\infty} c_{I,j}\sigma_{I,j}\phi_{I,j}(s), \\ \frac{\partial w_{II}}{\partial n}&=-\sum_{i=1}^{l-1} c_i\chi_{II,i}u_i(s)-\sum_{j=v \text{ or } v+1}^{\infty} c_{II,j}\sigma_{II,j}\phi_{II,j}(s), \end{aligned} \tag{42}$$

where σ_j are functions of λ but not of position, just as are c_i and c_j. Since $u_i(x, y)$ are eigenfunctions, $\partial u_i/\partial n$ must be continuous on γ, so $\chi_{I,i}=-\chi_{II,i}$.

With such expansions for $w(s)$, $\partial w/\partial n$ and $u_i(s)$ at hand, we proceed to find what conditions $w(s)$ must be subject to in order that the orthogonality conditions may be satisfied.

By Green's theorem, we have

$$\begin{aligned} &\int_{\Gamma_I+\gamma} p\left(w\frac{\partial u_i}{\partial n}-u_i\frac{\partial w}{\partial n}\right)ds \\ &=\iint_{R_I}\{wL[u_i]-u_iL[w]\}dxdy, \quad (i=1, 2, \cdots, l-1) \end{aligned}$$

and similar equations for R_{II}. Combining the equations for I and II together and putting $L[u_i]+\lambda_i\rho u_i=0$ and $L[w]+\lambda\rho w=0$, we obtain

$$\int_{I+II} p\left(w\frac{\partial u_i}{\partial n}-u_i\frac{\partial w}{\partial n}\right)ds=-(\lambda_i-\lambda)\iint_R \rho wu_idxdy.$$

The latter equation is further reduced to

$$\int_{\gamma_I+\gamma_{II}} p\left(w\frac{\partial u_i}{\partial n}-u_i\frac{\partial w}{\partial n}\right)ds=-(\lambda_i-\lambda)\int_R\int \rho w u_i dxdy,$$

because w is to satisfy the appropriate conditions, i.e., $\partial w/\partial n+\sigma w=0$ on Γ. So we must have

$$\int_{\gamma_I+\gamma_{II}} p\left(w\frac{\partial u_i}{\partial n}-u_i\frac{\partial w}{\partial n}\right)ds=0$$

in order that w will be orthogonal to u_i. Furthermore, since either w or $\partial w/\partial n$ is continuous on γ, we must have both

$$\int_{\gamma_I+\gamma_{II}} pw\frac{\partial u_i}{\partial n}ds=0 \quad \text{and} \quad \int_{\gamma_I+\gamma_{II}} pu_i\frac{\partial w}{\partial n}ds=0. \quad (43)$$

Substituting (40) and (42) into (43), we obtain

$$\int_{\gamma_I+\gamma_{II}} pw\frac{\partial u_i}{\partial n}ds=-\int_\gamma p\chi_{I,i}u_i(s)$$
$$\times\{\sum_{j=\nu+1}^{\infty} c_{I,j}\phi_{I,j}(s)\}ds+\text{similar term for } II=0 \quad (44)$$

and

$$\int_{\gamma_I+\gamma_{II}} pu_i\frac{\partial w}{\partial n}ds=-\int_\gamma pu_i(s)$$
$$\times\{\sum_{j=\nu+1}^{\infty} c_{I,j}\sigma_{I,j}\phi_{I,j}(s)\}ds+\text{similar term for } II=0. \quad (45)$$

Now we calculate

$$J(\lambda)=\int_{\gamma_I+\gamma_{II}} pw\frac{\partial w}{\partial n}ds$$
$$=-\sum_{i=1}^{l-1}c_i\int_\gamma pu_i(s)\{\sum_{j=1}^{l-1}c_j\chi_{I,j}u_j(s)\}ds-\sum_{j=\nu+1}^{\infty}(c_{I,j})^2\sigma_{I,j}$$
$$-\sum_{j=\nu+1}^{\infty}c_{I,j}\sigma_{I,j}\int_\gamma p\{\sum_{i=1}^{l-1}c_iu_i(s)\}\phi_{I,j}(s)ds$$
$$-\sum_{j=\nu+1}^{\infty}c_{I,j}\int_\gamma p\{\sum_{i=1}^{l-1}c_i\chi_{I,i}u_i(s)\}\phi_{I,j}(s)ds$$
$$+\text{similar terms for } II.$$

By virtue of (44) and (45) and $\chi_{I,j}=-\chi_{II,j}$, the above equation is reduced to

$$J(\lambda)=-\sum_{j=\nu+1}^{\infty}(c_{I,j})^2\sigma_{I,j}-\sum_{j=\nu \text{ or } \nu+1}^{\infty}(c_{II,j})^2\sigma_{II,j}. \quad (46)$$

Thus, if w is orthogonal to $u_i(i=1, 2, \cdots, l-1)$, the series expansion for $J(\lambda)$ will contain none of the first $(l-1)$ terms. The converse statement, however, is in general not true.

Let us rewrite (40) as

$$w_I(s)=\sum_1^{l-1}c_iu_i(s)+f_I(s)$$
$$w_{II}(s)=\sum_1^{l-1}c_iu_i(s)+f_{II}(s) \quad (40')$$

and (42) as

$$\frac{\partial w_I}{\partial n}=-\sum_1^{l-1}c_i\chi_{I,i}u_i(s)+\frac{\partial f_I}{\partial n}$$
$$\frac{\partial w_{II}}{\partial n}=-\sum_1^{l-1}c_i\chi_{II,i}u_i(s)+\frac{\partial f_{II}}{\partial n}. \quad (42')$$

If $f(s)$ or $\partial f/\partial n$ happens to be of suitable form, w will be orthogonal to $u_i(i=1, 2, \cdots, l-1)$; otherwise not. The function $w(s)$, which approximates $u_l(s)$, can be exactly right only when $J(\lambda)$ has no first $(l-1)$ terms.[16] Thus if we use $f(s)$ in place of $w(s)$ as the trial function, we will be able to get to the exact form of $J(\lambda)$, i.e.,

$$\int_{\gamma_I+\gamma_{II}} pu_l\frac{\partial u_l}{\partial n}ds\equiv 0,$$

by varying $f(s)$. It is to be noted that $f(s)$ does not approximate $u_l(s)$ but the function

$$u_l(s)-\sum_1^{l-1}c_iu_i(s).$$

Let us rewrite the expression for $J(\lambda)$ also. If we write $J(\lambda)$ for the upper-bound case, where w is matched

[16] For the case of perturbing boundary conditions, the equation $J(\lambda)=-\int_\Gamma p\sigma w^2ds$ or $J(\lambda)=-\int_\Gamma p/\sigma(\partial w/\partial n)^2ds$ will be reduced to

$$-\sum_{j=l}^{\infty}\sigma_jc_j^2=-\int_{\Gamma'}p\sigma(\sum_{j=l}^{\infty}c_j\phi_j)^2ds$$

or

$$-\sum_{j=l}^{\infty}\sigma_jc_j^2=-\int_{\Gamma'}p/\sigma(\sum_{j=l}^{\infty}c_j\sigma_j\phi_j)^2ds,$$

Γ' being that part of Γ where the boundary condition is perturbed and a complete set of appropriate orthogonal functions ϕ_j exist. Thus

$$J(\lambda)+\int_\Gamma p\sigma w^2ds=\int_{\Gamma'}p\sigma f^2ds-\sum_{j=l}^{\infty}\sigma_j\left(\int_{\Gamma'}pf\phi_jds\right)^2$$

or

$$J(\lambda)+\int_\Gamma p/\sigma(\partial w/\partial n)^2ds=\int_{\Gamma'}p/\sigma(\partial f/\partial n)^2ds$$
$$-\sum_{j=l}^{\infty}1/\sigma_j\left(\int_{\Gamma'}p(\partial f/\partial n)\phi_jds\right)^2$$

will have no first $(l-1)$ terms if w is orthogonal to u_i $(i=1, 2, \cdots, l-1)$. The rest of the discussion will be similar to that for the case of perturbing continuity conditions.

on γ, as

$$J_+(\lambda)=-\sum_1^\infty \sigma_{I,j}(\lambda)\left(\int_\gamma pw\phi_{I,j}ds\right)^2 - \sum_1^\infty \sigma_{II,j}(\lambda)\left(\int_\gamma pw\phi_{II,j}ds\right)^2, \quad (47)$$

the corresponding expression for the lower-bound case, where $\partial w/\partial n$ is matched, will be

$$J_-(\lambda)=-\sum_1^\infty \frac{1}{\sigma_{I,j}(\lambda)}\left(\int_\gamma p\cdot\frac{\partial w}{\partial n}\cdot\phi_{I,j}ds\right)^2 - \sum_1^\infty \frac{1}{\sigma_{II,j}(\lambda)}\left(\int_\gamma p\cdot\frac{\partial w}{\partial n}\cdot\phi_{II,j}ds\right)^2. \quad (48)$$

It is obvious that the corresponding terms of $J_+(\lambda)$ and $J_-(\lambda)$ always have the same sign. For the fundamental mode there is no question about the orthogonality conditions, so only the first term can have a sign different from all the rest. Otherwise we would get at least two values for λ. This could not be possible if w has not zeros inside R. As λ increases, the higher terms will change sign. The second term will have changed sign for the second mode and so the lth term changed for the lth mode. As long as the first $l-1$ terms in $J(\lambda)$ are missing, there can be only one term which has different sign from all the rest. As shown by Schwinger such forms of $J(\lambda)$ are indeed definite for a certain fixed value of λ, being λ_l in the present case. Thus, if we can show that $\partial J_+(\lambda)/\partial\lambda$ has definite sign and the sign is always opposite to that of $\partial J_-(\lambda)/\partial\lambda$, then the value of λ calculated by solving $J_+=0$ or $J_-=0$ will always be greater or lower than the true value.

Though the above result may be anticipated from the proof we had for (34), it is desirable to determine how each term of the series varies with λ. We consider a single coefficient $\sigma_j(\lambda)$ used in (47) and calculate $\partial/\partial\lambda\sigma_j(\lambda)$. Since σ_j satisfies $\partial\phi_j/\partial n+\sigma_j\phi_j=0$ on γ by definition, so

$$\int_\gamma p\phi_j\frac{\partial\phi_j}{\partial n}ds=-\int_\gamma p\sigma_j\phi^2ds.$$

Since $\phi_j(x, y)$ satisfies the differential equation, we have by Green's theorem

$$\iint_{I\text{ or }II}\{p(\nabla\phi_j)^2+q\phi_j^2\}dxdy + \iint_{I\text{ or }II}\phi_jL[\phi_j]dxdy=\int_{I\text{ or }II}p\phi_j\frac{\partial\phi_j}{\partial n}ds$$

i.e.,

$$D_{I\text{ or }II}[\phi_j, \phi_j]-\lambda\iint_{I\text{ or }II}\rho\phi_j^2dxdy=\int_{I\text{ or }II}p\phi\frac{\partial\phi_j}{j\partial n}ds.$$

Differentiating with respect to λ, we obtain

$$\frac{\partial}{\partial\lambda}D_{I\text{ or }II}[\phi_j, \phi_j]-\iint_{I\text{ or }II}\rho\phi_j^2dxdy - 2\lambda\iint_{I\text{ or }II}\rho\phi_j\frac{\partial\phi_j}{\partial\lambda}dxdy=\frac{\partial}{\partial\lambda}\int_{I\text{ or }II}p\phi_j\frac{\partial\phi_j}{\partial n}ds.$$

Since $\phi_j(x, y)$ as $w(x, y)$ satisfies the appropriate boundary conditions on Γ, i.e., $\partial\phi_j/\partial n+\sigma\phi_j=0$ on Γ_I and Γ_{II}, the right-hand side of the above equation becomes

$$\frac{\partial}{\partial\lambda}\int_{I\text{ or }II}p\phi_j\frac{\partial\phi_j}{\partial n}ds = -\frac{\partial}{\partial\lambda}\int_{\Gamma_I\text{ or }\Gamma_{II}}p\sigma\phi_j^2ds-\frac{\partial}{\partial\lambda}\int_\gamma p\sigma_j\phi_j^2ds$$

while the left-hand side is

$$2D_{I\text{ or }II}\left[\phi_j, \frac{\partial\phi_j}{\partial\lambda}\right]-\iint_{I\text{ or }II}\rho\phi_j^2dxdy-2\lambda\iint_{I\text{ or }II}\rho\phi_j\frac{\partial\phi_j}{\partial\lambda}dxdy$$

$$=2\int_{I\text{ or }II}p\frac{\partial\phi_j}{\partial\lambda}\cdot\frac{\partial\phi_j}{\partial n}ds-\iint_{I\text{ or }II}\rho\phi_j^2dxdy$$

$$=-\frac{\partial}{\partial\lambda}\int_{\Gamma_I\text{ or }\Gamma_{II}}p\sigma\phi_j^2ds-\frac{\partial}{\partial\lambda}\int_\gamma p\sigma_j\phi_j^2ds + \frac{\partial\sigma_j}{\partial\lambda}\int_\gamma p\phi_j^2ds-\iint_{I\text{ or }II}\rho\phi_j^2dxdy.$$

In the last step of derivation, we have made use of the fact that σ is the given boundary condition, and so independent of λ, while σ_j depends on λ but not on position.

Therefore by equating the two sides we obtain

$$\frac{\partial\sigma_j}{\partial\lambda}\int_\gamma p\phi_j^2ds=\iint_{I\text{ or }II}\rho\phi_j^2dxdy.$$

Since $p>0$ and $\rho>0$, so

$$\partial\sigma_j/\partial\lambda>0,$$

so also

$$(\partial/\partial\lambda)(1/\sigma_j)<0.$$

Thus we obtain

$$\partial J_+(\lambda)/\partial\lambda<0 \quad (49)$$

while

$$[\partial J_-(\lambda)]/\partial\lambda>0, \quad (50)$$

because every term of J_+ or J_- changes in the same sense.

Now we have proved that, by dropping out the first $l-1$ terms in J_+ or J_- the error in the calculated value of λ will always have one sign. In fact we can be spared the above proof by utilizing the maximum-minimum property of eigenvalues. According to this property the sense of error for the lower-bound calculation cannot be changed by using a trial function which is not exactly orthogonal to the $(l-1)$ u_i's. Since J_- is positive definite, $\partial J_-/\partial\lambda$ must be positive also. By the reciprocal nature of the role played by σ_j in J_+ and J_- we must have $\partial J_+/\partial\lambda<0$. Thus we obtain the same result in a more elegant way.

The original requirements, that $w(x, y)$ should be orthogonal to all the lower eigenfunctions, have been reduced to the requirements that $f(s)$, which approximates

$$u_l(s)-\sum_1^{l-1} c_i u_i(s),$$

should be orthogonal to the first $l-1$ known functions ϕ_j's. This not only simplifies the process but also makes the upper-bound calculation possible without knowing the true eigenfunctions.

Using such a trial function we are assured that the calculated value of λ will always lie above or always lie below λ_l but not always very close. In any case we must use good enough trial functions. The higher terms of the series will be determined by the singular character of the function while the lower ones by the orthogonality requirements. Thus $f(s)$ (and so $w(s)$) can actually be chosen to be sufficiently near to the true function.

We may further improve the accuracy by successive approximations. For example, we take

$$f'(s)=f(s)+\beta\sum_{v+1}^{2v+1 \text{ or } 2v+2}(c_{I,j}\phi_{I,j}(s)),$$

where

$$\sum_{v+1}^{2v+1 \text{ or } 2v+2} c_{I,j}\phi_{I,j}(s)$$

is to be orthogonal to $\phi_{II,i}(i=1, 2, \cdots, v \text{ or } v+1)$, so $c_{I,j}(j=v+1, v+2, \cdots, 2v+1 \text{ or } 2v+2)$ are known to a constant factor. Apply variation on $J(\lambda)$ with respect to β, i.e., put $\delta_\beta J(\lambda)=0$. This determines β and so all the coefficients $\beta c_{I,j}$. The value calculated by using $f'(s)$ will be more accurate than by using $f(s)$. This process can be carried on successively to whatever accuracy we may desire. In most practical cases, however, the trial function may be chosen to be sufficiently accurate, so there is seldom need to go further than a second approximation.

The author wishes to express his deep gratitude to the late Professor W. W. Hansen, from whom he received invaluable advice and friendly encouragement.

Letters to the Editor

Single Crystal X-Ray Techniques for Use with Alloys Requiring Protective Atmospheres*

JOSEPH SINGER AND F. H. ELLINGER
Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico
January 27, 1950

PROCEDURES similar to those to be described undoubtedly have been used by others; it was felt, however, that their accessibility in the literature might be desirable.

Our procedures are complicated by the need for handling our alloys in protective atmospheres. An ordinary type of "dry box" equipped with rubber gloves is the preparation chamber through which the inert gas constantly flows. Single crystal fragments of intermetallic compounds have been obtained by shattering ordinary ingots with a hammer. The fragments are viewed in a glass dish against a background of dull black paper. With the aid of suitable tweezers a glass fiber is touched to thin shellac and brought up to a selected fragment. With ingots of small grain size, the fragment selected should be as small as can be seen, preferably with the aid of magnification. If now the mounted fragment were to be dropped into a capillary tube of the type shown in Fig. 1, in which the evacuation and sealing-off is to be done, the specimen would very likely be lost. The following are two ways of controlling the placing of the specimen inside the capillary tube.

One procedure is to place the mounting fiber into a shorter, thin-walled plastic tube which has been formed around a wire smaller than the glass capillary. The metallic fragment is set barely to protrude beyond the plastic and the back end of its mount is fastened to the plastic by clamping. This assemblage can be dropped safely into the glass tube. Figure 1 illustrates this procedure.

A second method consists of using a long fiber, drawn fine only at the end where the specimen is to be attached. Each such fiber is made the right length for a particular capillary tube so that, when the specimen is at a desired position, the back end of the fiber is at the mouth of the tube. With the specimen suitably situated, the tube is held horizontally and shellac brushed on the back end where the fiber touches the wall at the entrance to the tube. When the shellac is dry, the assembly can be held vertically for the evacuation and sealing-off without displacing the specimen. The sealing-off will fasten the fiber within the capillary so that the sample is held rigidly at a desired position.

It is frequently necessary to mount a specimen about a second axis, once a set of single crystal data has been obtained for one

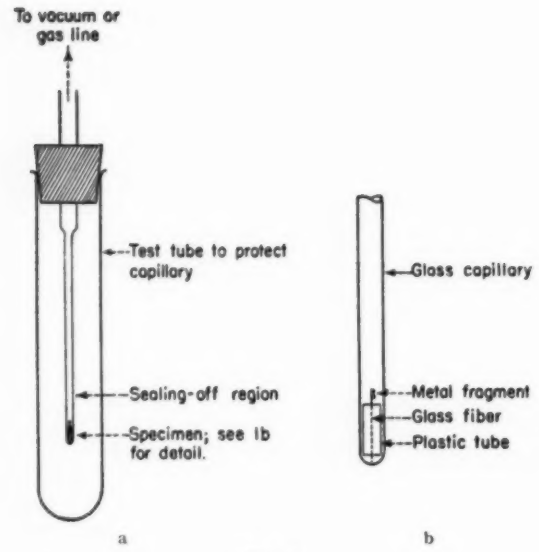


FIG. 1.

axis. With fragments of alloys the usual crystallographic aids for locating an axial direction are not available. However, the position of an axis can be deduced from the previous data. This direction can be set parallel to the x-ray beam. A long, rather heavy glass fiber is set up on a lump of putty-like material and worked over towards the crystal mount until the two pieces of glass touch and the long fiber parallels the selected direction. A small quantity of shellac is deposited around the intersection and is allowed to harden thoroughly. Finally the old mount is carefully detached from the goniometer head, the new fiber shortened, and the crystal aligned on its new axis.

With a little practice one can perform the dry box operations skillfully despite the intervening rubber gloves. The effort is amply repaid with the type of x-ray diffraction data resulting.

* This letter is based on work performed at Los Alamos Scientific Laboratory of the University of California under contract W-7405-Eng-36.

Errata: On the Diffraction of an Electromagnetic Wave through a Plane Screen

[J. App. Phys. 20, 760 (1949)]

J. W. MILES
Department of Engineering, University of California, Los Angeles, California

THE scattering cross section for plane, electromagnetic waves normally incident upon a circular aperture in an infinite, plane, perfectly conducting screen given by us in a recent paper[1] was calculated erroneously. This error, in turn, was due to an error

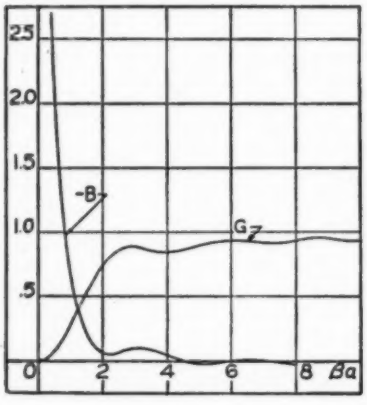


FIG. 9. The conductance and susceptance of a circular aperture for normal incidence of a plane, electromagnetic wave.

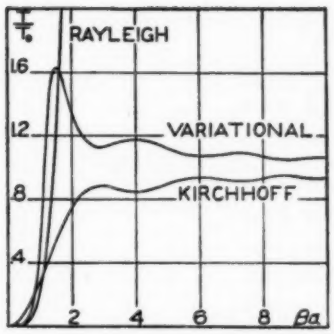


FIG. 10. Scattering cross section (relative to geometric) of a circular aperture for normal incidence of a plane, electromagnetic wave.

in the computation of the aperture susceptance (B). The results have been recalculated, and Figs. 9 and 10 in the original paper should be replaced by those given here.

The error in our calculations was brought to our attention by Professor Harold Levine (Harvard University), who has carried out similar (variational) calculations,[2] using as first and second approximations to the aperture field $[1-(r/a)^2]^{\frac{1}{2}}$ and $A_1[1-(r/a)^2]^{\frac{1}{2}}+A_2[1-(r/a)^2]^{\frac{3}{2}}$, respectively. Our calculation, which uses the static approximation $[1-(r/a)^2]^{\frac{1}{2}}$ for the leading term for small βa and the incident field for the remaining terms, does not predict the first maximum of the scattering cross section as accurately as that of Levine. (We obtain $(T/T_0)_{max}=1.63$ at $\beta a=1.5$; Levine obtains $(T/T_0)_{max}=1.78$ at $\beta a=1.4$.) But it is somewhat more accurate (relative to Levine's first approximation) for large βa, all of which is as should be expected from the nature of the field approximations.

Two other errors in the paper should be pointed out. Equation (86b) should have read

$$B_e(\beta a)=\sum_{m=0}^{\infty}\frac{(-)^m(\beta a)^{2m+1}}{(2m+2)[\Gamma(m+\frac{3}{2})]^2}-(3\pi/8\beta a). \qquad (86b)$$

(This error was typographical and not connected with the calculations.) Moreover, contrary to our statement, Magnus[3] *did* solve the integral equation which he formulated for the Sommerfeld problem,[3] and we should have stated merely that his solution was not as direct as the Wiener-Hopf approach.

[1] J. App. Phys. 20, 760–771 (1949).
[2] Progress Report 6, Cruft Laboratory, Harvard University, Contract N5-OR1-76 (January 1, 1948).
[3] W. Magnus, Zeits. f. Physik 117, 168 (1941). See also W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Chelsea Publishing Company, New York, 1949), pp. 142, 143.

Erratum: Distribution of Phases in Two-Phase Solids

[J. App. Phys. 20, 1206–1208 (1949)]

LEONARD D. JAFFE
Watertown Arsenal Laboratory, Watertown, Massachusetts

AN error in proof-reading made the inequality on page 1208, line 15, inconsistent with the rest of the paper. It should read $r_1<7\times10^{-6}$ cm.

New Books

Letter Symbols for Aeronautical Sciences

BY SEVEN SPONSORING OR COLLABORATING ORGANIZATIONS. Pp. 20. Published by American Standards Association, New York, 1950. Price $1.25.